Conjunctive Query Answering via a Fragment of Set Theory (Extended Version)

Domenico Cantone, Marianna Nicolosi-Asmundo, and
Daniele Francesco Santamaria

University of Catania, Dept. of Mathematics and Computer Science
email: {cantone,nicolosi,santamaria}@dmi.unict.it

Abstract. We address the problem of Conjunctive Query Answering (CQA) for the description logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R},\times}\rangle(\mathbf{D})$ ($\mathcal{DL}_{\mathbf{D}}^{4,\times}$, for short) which extends the logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R}}\rangle(\mathbf{D})$ with Boolean operations on concrete roles and with the product of concepts.
The result is obtained by formalizing $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge bases and $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive queries in terms of formulae of the four-level set-theoretic fragment $\mathsf{4LQS}^{\mathsf{R}}$, which admits a restricted form of quantification on variables of the first three levels and on pair terms. We solve the CQA problem for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ through a decision procedure for the satisfiability problem of $\mathsf{4LQS}^{\mathsf{R}}$. We further define a KE-tableau based procedure for the same problem, more suitable for implementation purposes, and analyze its computational complexity.

1 Introduction

In the last few years, results from Computable Set Theory have been used as a means to represent and reason about description logics and rule languages for the semantic web. For instance, in [6,8,9], fragments of set theory with constructs related to *multi-valued maps* have been studied and applied to the realm of knowledge representation. In [10], an expressive description logic, called $\mathcal{DL}\langle\mathsf{MLSS}_{2,m}^{\times}\rangle$, has been introduced and the consistency problem for $\mathcal{DL}\langle\mathsf{MLSS}_{2,m}^{\times}\rangle$-knowledge bases has been proved **NP**-complete. The description logic $\mathcal{DL}\langle\mathsf{MLSS}_{2,m}^{\times}\rangle$ has been extended with additional constructs and SWRL rules in [9], proving that the decision problem for the resulting logic, called $\mathcal{DL}\langle\forall_{\mathbf{0,2}}^{\pi}\rangle$, is still **NP**-complete under suitable conditions. The description logic $\mathcal{DL}\langle\forall_{\mathbf{0,2}}^{\pi}\rangle$ has been extended with some *metamodelling* features in [6]. In [7], the description logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R}}\rangle(\mathbf{D})$ (more simply referred to as $\mathcal{DL}_{\mathbf{D}}^{4}$) has been introduced. $\mathcal{DL}_{\mathbf{D}}^{4}$ can be represented in the decidable four-level stratified fragment of set theory $\mathsf{4LQS}^{\mathsf{R}}$ involving a restricted form of quantification over variables of the first three levels and pair terms (cf. [4]). The logic $\mathcal{DL}_{\mathbf{D}}^{4}$ admits concept constructs such as full negation, union and intersection of concepts, concept domain and range, existential quantification and min cardinality on the left-hand side of inclusion axioms. It also supports role constructs such as role chains on the left hand side of inclusion axioms, union, intersection, and complement of abstract roles, and

properties on roles such as transitivity, symmetry, reflexivity, and irreflexivity. It admits datatypes, a simple form of concrete domains that are relevant in real world applications. The consistency problem for $\mathcal{DL}^4_{\mathbf{D}}$-knowledge bases has been proved decidable in [7] by means of a reduction to the satisfiability problem for $\mathsf{4LQS^R}$, proved decidable in [4]. It has also been proved, under not very restrictive constraints, that the consistency problem for $\mathcal{DL}^4_{\mathbf{D}}$-knowledge bases is **NP**-complete. Finally, we mention that the papers [6–10] are concerned with traditional research issues for description logics mainly focused on the parts of a knowledge base representing conceptual information, namely the TBox and the RBox, where the principal reasoning services are subsumption and satisfiability.

In this paper we exploit decidability results presented in [4, 7] to deal with reasoning services for knowledge bases involving ABoxes. The most basic service to query the instance data is *instance retrieval*, i.e., the task of retrieving all individuals that instantiate a class C, and, dually, all named classes C that an individual belongs to. In particular, a powerful way to query ABoxes is the *Conjunctive Query Answering* task (CQA). CQA is relevant in the context of description logics and, in particular, for real world applications based on semantic web technologies, since it provides a mechanism allowing users and applications to interact with ontologies and data. The task of CQA has been studied for several well-known description logics (cf. [1–3, 13–18, 20–23]). In particular, we introduce the description logic $\mathcal{DL}\langle\mathsf{4LQS^{R,\times}}\rangle(\mathbf{D})$ ($\mathcal{DL}^{4,\times}_{\mathbf{D}}$, for short), extending $\mathcal{DL}^4_{\mathbf{D}}$ with Boolean operations on concrete roles and with the product of concepts. Then we define the CQA problem for $\mathcal{DL}^{4,\times}_{\mathbf{D}}$ and prove its decidability via a reduction to the CQA problem for $\mathsf{4LQS^R}$, whose decidability follows from that of the satisfiability problem for $\mathsf{4LQS^R}$ (proved in [4]). Finally, we present a KE-tableau based procedure that, given a $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-query Q and a $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-knowledge base $\mathcal{KB}$ represented in set-theoretic terms, determines the answer set of Q with respect to $\mathcal{KB}$, providing also some complexity results. The choice of the KE-tableau system [11] is motivated by the fact that this variant of the tableau method allows one to construct trees whose distinct branches define mutually exclusive situations thus avoiding the proliferation of redundant branches, typical of semantic tableaux.

2 Preliminaries

2.1 The set-theoretic fragment $\mathsf{4LQS^R}$

It is convenient to first introduce the syntax and semantics of a more general four-level quantified language, denoted $\mathsf{4LQS}$. Then we provide some restrictions on quantified formulae of $\mathsf{4LQS}$ that characterize $\mathsf{4LQS^R}$. We recall that the satisfiability problem for $\mathsf{4LQS^R}$ has been proved decidable in [4].

$\mathsf{4LQS}$ involves four collections, $\mathcal{V}_i$, of variables of sort i, for $i = 0, 1, 2, 3$. Variables of sort i, for $i = 0, 1, 2, 3$, will be denoted by $X^i, Y^i, Z^i, \ldots$ (in particular, variables of sort 0 will also be denoted by $x, y, z, \ldots$). In addition to variables, $\mathsf{4LQS}$ involves also *pair terms* of the form $\langle x, y\rangle$, with $x, y \in \mathcal{V}_0$.

$\mathsf{4LQS}$-*quantifier-free atomic formulae* are classified as:

- level 0: $x = y$, $x \in X^1$, $\langle x, y\rangle = X^2$, $\langle x, y\rangle \in X^3$;
- level 1: $X^1 = Y^1$, $X^1 \in X^2$;
- level 2: $X^2 = Y^2$, $X^2 \in X^3$.

$\mathsf{4LQS}$ *purely universal formulae* are classified as:

- level 1: $(\forall z_1)\ldots(\forall z_n)\varphi_0$, where $z_1,\ldots,z_n \in \mathcal{V}_0$ and φ_0 is any propositional combination of quantifier-free atomic formulae of level 0;
- level 2: $(\forall Z_1^1)\ldots(\forall Z_m^1)\varphi_1$, where $Z_1^1,\ldots,Z_m^1 \in \mathcal{V}_1$ and φ_1 is any propositional combination of quantifier-free atomic formulae of levels 0 and 1, and of purely universal formulae of level 1;
- level 3: $(\forall Z_1^2)\ldots(\forall Z_p^2)\varphi_2$, where $Z_1^2,\ldots,Z_p^2 \in \mathcal{V}_2$ and φ_2 is any propositional combination of quantifier-free atomic formulae and of purely universal formulae of levels 1 and 2.

$\mathsf{4LQS}$-formulae are all the propositional combinations of quantifier-free atomic formulae of levels 0, 1, 2 and of purely universal formulae of levels 1, 2, 3.

Let φ be a $\mathsf{4LQS}$-formula. Without loss of generality, we can assume that φ contains only $\neg$, $\wedge$, $\vee$ as propositional connectives. Further, let S_φ be the syntax tree for a $\mathsf{4LQS}$-formula φ,[1] and let ν be a node of S_φ. We say that a $\mathsf{4LQS}$-formula ψ occurs within φ at position ν if the subtree of S_φ rooted at ν is identical to S_ψ. In this case we refer to ν as an occurrence of ψ in φ and to the path from the root of S_φ to ν as its occurrence path. An occurrence of ψ within φ is *positive* if its occurrence path deprived by its last node contains an even number of nodes labelled by a $\mathsf{4LQS}$-formula of type $\neg\chi$. Otherwise, the occurrence is said to be *negative*.

The variables $z_1,\ldots,z_n$ are said to occur *quantified* in $(\forall z_1)\ldots(\forall z_n)\varphi_0$. Likewise, $Z_1^1,\ldots,Z_m^1$ and $Z_1^2,\ldots,Z_p^2$ occur quantified in $(\forall Z_1^1)\ldots(\forall Z_m^1)\varphi_1$ and in $(\forall Z_1^2)\ldots(\forall Z_p^2)\varphi_2$, respectively. A variable occurs *free* in a $\mathsf{4LQS}$-formula φ if it does not occur quantified in any subformula of φ. For $i = 0,1,2,3$, we denote with $\mathtt{Var}_i(\varphi)$ the collections of variables of level i occurring free in φ.

A (level 0) substitution $\sigma \coloneqq \{x_1/y_1,\ldots,x_n/y_n\}$ is the mapping $\varphi \mapsto \varphi\sigma$ such that, for any given $\mathsf{4LQS}$-formula φ, $\varphi\sigma$ is the $\mathsf{4LQS}$-formula obtained from φ by replacing the free occurrences of the variables $x_1,\ldots,x_n$ in φ with the variables $y_1,\ldots,y_n$, respectively. We say that a substitution σ is free for φ if the formulae φ and $\varphi\sigma$ have exactly the same occurrences of quantified variables.

A $\mathsf{4LQS}$-*interpretation* is a pair $\boldsymbol{\mathcal{M}} = (D, M)$ where D is a non-empty collection of objects (called *domain* or *universe* of $\boldsymbol{\mathcal{M}}$) and M is an assignment over the variables in $\mathcal{V}_i$, for $i = 0,1,2,3$, such that:

$$MX^0 \in D, \quad MX^1 \in \mathcal{P}(D), \quad MX^2 \in \mathcal{P}(\mathcal{P}(D)), \quad MX^3 \in \mathcal{P}(\mathcal{P}(\mathcal{P}(D))),$$

where $X^i \in \mathcal{V}_i$, for $i = 0,1,2,3$, and $\mathcal{P}(s)$ denotes the powerset of s.

Pair terms are interpreted *à la* Kuratowski, and therefore we put

[1] The notion of syntax tree for $\mathsf{4LQS}$-formulae is similar to the notion of syntax tree for formulae of first-order logic. A precise definition of the latter can be found in [12].

$$M\langle x, y\rangle \coloneqq \{\{Mx\}, \{Mx, My\}\}.$$

Next, let

- $\boldsymbol{\mathcal{M}} = (D, M)$ be a $\mathsf{4LQS}$-interpretation,
- $x_1, \ldots, x_n \in \mathcal{V}_0, \quad X_1^1, \ldots, X_m^1 \in \mathcal{V}_1, \quad X_1^2, \ldots, X_p^2 \in \mathcal{V}_2$, and
- $u_1, \ldots, u_n \in D, \quad U_1^1, \ldots, U_m^1 \in \mathcal{P}(D), \quad U_1^2, \ldots, U_p^2 \in \mathcal{P}(\mathcal{P}(D))$.

By $\boldsymbol{\mathcal{M}}[\boldsymbol{x}/\boldsymbol{u}, \boldsymbol{X}^1/\boldsymbol{U}^1, \boldsymbol{X}^2/\boldsymbol{U}^2]$, we denote the interpretation $\boldsymbol{\mathcal{M}'} = (D, M')$ such that $M'x_i = u_i$ (for $i = 1, \ldots, n$), $M'X_j^1 = U_j^1$ (for $j = 1, \ldots, m$), $M'X_k^2 = U_k^2$ (for $k = 1, \ldots, p$), and which otherwise coincides with M on all remaining variables. For a $\mathsf{4LQS}$-interpretation $\boldsymbol{\mathcal{M}} = (D, M)$ and a $\mathsf{4LQS}$-formula φ, the satisfiability relationship $\boldsymbol{\mathcal{M}} \models \varphi$ is defined inductively over the structure of φ as follows. Quantifier-free atomic formulae are evaluated in a standard way according to the usual meaning of the predicates '$\in$' and '$=$', and purely universal formulae are evaluated as follows:

- $\boldsymbol{\mathcal{M}} \models (\forall z_1) \ldots (\forall z_n)\varphi_0$ iff $\boldsymbol{\mathcal{M}}[\boldsymbol{z}/\boldsymbol{u}] \models \varphi_0$, for all $\boldsymbol{u} \in D^n$;
- $\boldsymbol{\mathcal{M}} \models (\forall Z_1^1) \ldots (\forall Z_m^1)\varphi_1$ iff $\boldsymbol{\mathcal{M}}[\boldsymbol{Z}^1/\boldsymbol{U}^1] \models \varphi_1$, for all $\boldsymbol{U}^1 \in \big(\mathcal{P}(D)\big)^m$;
- $\boldsymbol{\mathcal{M}} \models (\forall Z_1^2) \ldots (\forall Z_p^2)\varphi_2$ iff $\boldsymbol{\mathcal{M}}[\boldsymbol{Z}^2/\boldsymbol{U}^2] \models \varphi_2$, for all $\boldsymbol{U}^2 \in \big(\mathcal{P}(\mathcal{P}(D))\big)^p$.

Finally, compound formulae are interpreted according to the standard rules of propositional logic. If $\boldsymbol{\mathcal{M}} \models \varphi$, then $\boldsymbol{\mathcal{M}}$ is said to be a $\mathsf{4LQS}$-model for φ. A $\mathsf{4LQS}$-formula is said to be *satisfiable* if it has a $\mathsf{4LQS}$-model. A $\mathsf{4LQS}$-formula is *valid* if it is satisfied by all $\mathsf{4LQS}$-interpretations.

We are now ready to present the fragment $\mathsf{4LQS^R}$ of $\mathsf{4LQS}$ of our interest. This is the collection of the formulae ψ of $\mathsf{4LQS}$ fulfilling the restrictions:

1. for every purely universal formula $(\forall Z_1^1) \ldots (\forall Z_m^1)\varphi_1$ of level 2 occurring in ψ and every purely universal formula $(\forall z_1) \ldots (\forall z_n)\varphi_0$ of level 1 occurring negatively in φ_1, φ_0 is a propositional combination of quantifier-free atomic formulae of level 0 and the condition
$$\neg\varphi_0 \rightarrow \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m} z_i \in Z_j^1$$
is a valid $\mathsf{4LQS}$-formula (in this case we say that $(\forall z_1) \ldots (\forall z_n)\varphi_0$ is *linked to the variables* $Z_1^1, \ldots, Z_m^1$);
2. for every purely universal formula $(\forall Z_1^2) \ldots (\forall Z_p^2)\varphi_2$ of level 3 in ψ:
   - every purely universal formula of level 1 occurring negatively in φ_2 and not occurring in a purely universal formula of level 2 is only allowed to be of the form
   $$(\forall z_1) \ldots (\forall z_n)\neg(\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{n} \langle z_i, z_j\rangle = Y_{ij}^2),$$
   with $Y_{ij}^2 \in \mathcal{V}^2$, for $i, j = 1, \ldots, n$;
   - purely universal formulae $(\forall Z_1^1) \ldots (\forall Z_m^1)\varphi_1$ of level 2 may occur only positively in φ_2.

Restriction 1 has been introduced for technical reasons concerning the decidability of the satisfiability problem for the fragment, while restriction 2 allows

one to define binary relations and several operations on them (for space reasons details are not included here but can be found in [4]).

The semantics of $\mathsf{4LQS}^{\mathsf{R}}$ plainly coincides with that of $\mathsf{4LQS}$.

2.2 The logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R},\times}\rangle(\mathbf{D})$

The description logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R},\times}\rangle(\mathbf{D})$ (more simply referred to as $\mathcal{DL}_{\mathbf{D}}^{4,\times}$) is the extension of the logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R}}\rangle(\mathbf{D})$ (for short $\mathcal{DL}_{\mathbf{D}}^{4}$) presented in [7] in which Boolean operations on concrete roles and the product of concepts are admitted. Analogously to $\mathcal{DL}_{\mathbf{D}}^{4}$, the logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ supports concept constructs such as full negation, union and intersection of concepts, concept domain and range, existential quantification and min cardinality on the left-hand side of inclusion axioms, role constructs such as role chains on the left hand side of inclusion axioms, union, intersection, and complement of roles, and properties on roles such as transitivity, symmetry, reflexivity, and irreflexivity.

As far as the construction of role inclusion axioms is concerned, $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ is more liberal than $\mathcal{SROIQ}(\mathbf{D})$ (the logic underlying the most expressive Ontology Web Language 2 profile, OWL 2 DL [24]), since the roles involved are not required to be subject to any ordering relationship, and the notion of simple role is not needed. $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ treats derived datatypes by admitting datatype terms constructed from data ranges by means of a finite number of applications of the Boolean operators. Basic and derived datatypes can be used inside inclusion axioms involving concrete roles.

Datatypes are defined according to [19] as follows. Let $\mathbf{D} = (N_D, N_C, N_F, \cdot^{\mathbf{D}})$ be a *datatype map*, where N_D is a finite set of datatypes, N_C is a map assigning a set of constants $N_C(d)$ to each datatype $d \in N_D$, N_F is a map assigning a set of facets $N_F(d)$ to each $d \in N_D$, and $\cdot^{\mathbf{D}}$ is a map assigning

(i) a datatype interpretation $d^{\mathbf{D}}$ to each datatype $d \in N_D$,

(ii) a facet interpretation $f^{\mathbf{D}} \subseteq d^{\mathbf{D}}$ to each facet $f \in N_F(d)$, and

(iii) a data value $e_d^{\mathbf{D}} \in d^{\mathbf{D}}$ to every constant $e_d \in N_C(d)$.

We shall assume that the interpretations of the datatypes in N_D are nonempty pairwise disjoint sets.

A *facet expression* for a datatype $d \in N_D$ is a formula ψ_d constructed from the elements of $N_F(d) \cup \{\top_d, \bot_d\}$ by applying a finite number of times the connectives $\neg$, $\wedge$, and $\vee$. The function $\cdot^{\mathbf{D}}$ is extended to facet expressions for $d \in N_D$ by putting for $f, f_1, f_2 \in N_F(d)$

- $\top_d^{\mathbf{D}} = d^{\mathbf{D}}$,
- $\bot_d^{\mathbf{D}} = \emptyset$,
- $(\neg f)^{\mathbf{D}} = d^{\mathbf{D}} \setminus f^{\mathbf{D}}$,
- $(f_1 \wedge f_2)^{\mathbf{D}} = f_1^{\mathbf{D}} \cap f_2^{\mathbf{D}}$,
- $(f_1 \vee f_2)^{\mathbf{D}} = f_1^{\mathbf{D}} \cup f_2^{\mathbf{D}}$.

A *data range* dr for $\mathbf{D}$ is either a datatype $d \in N_D$, or a finite enumeration of datatype constants $\{e_{d_1}, \dots, e_{d_n}\}$, with $e_{d_i} \in N_C(d_i)$ and $d_i \in N_D$, or a facet expression ψ_d, for $d \in N_D$, or their complementation.

Let $\mathbf{R_A}$, $\mathbf{R_D}$, $\mathbf{C}$, $\mathbf{Ind}$ be denumerable pairwise disjoint sets of abstract role names, concrete role names, concept names, and individual names, respectively. We assume that the set of abstract role names $\mathbf{R_A}$ contains a name U denoting the universal role.
(a) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-datatype, (b) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concept, (c) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role, and (d) $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concrete role terms are constructed according to the following syntax rules:

(a) $t_1, t_2 \longrightarrow dr \mid \neg t_1 \mid t_1 \sqcap t_2 \mid t_1 \sqcup t_2 \mid \{e_d\}$,

(b) $C_1, C_2 \longrightarrow A \mid \top \mid \bot \mid \neg C_1 \mid C_1 \sqcup C_2 \mid C_1 \sqcap C_2 \mid \{a\} \mid \exists R.\mathit{Self} \mid \exists R.\{a\} \mid \exists P.\{e_d\}$,

(c) $R_1, R_2 \longrightarrow S \mid U \mid R_1^- \mid \neg R_1 \mid R_1 \sqcup R_2 \mid R_1 \sqcap R_2 \mid R_{C_1|} \mid R_{|C_1} \mid R_{C_1 \mid C_2} \mid id(C) \mid C_1 \times C_2$,

(d) $P_1, P_2 \longrightarrow T \mid \neg P_1 \mid P_1 \sqcup P_2 \mid P_1 \sqcap P_2 \mid P_{C_1|} \mid P_{|t_1} \mid P_{C_1|t_1}$,

where dr is a data range for $\mathbf{D}$, t_1, t_2 are data-type terms, e_d is a constant in $N_C(d)$, a is an individual name, A is a concept name, C_1, C_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concept terms, S is an abstract role name, R, R_1, R_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role terms, T is a concrete role name, and P, P_1, P_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concrete role terms.

A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base is a triple $\mathcal{KB} = (\mathcal{R}, \mathcal{T}, \mathcal{A})$ such that $\mathcal{R}$ is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*RBox*, $\mathcal{T}$ is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*TBox*, and $\mathcal{A}$ a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*ABox* (see next).

A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*RBox* is a collection of statements of the following forms:

$$
\begin{array}{ccccc}
R_1 \equiv R_2, & R_1 \sqsubseteq R_2, & R_1 \dots R_n \sqsubseteq R_{n+1}, & \mathsf{Sym}(R_1), & \mathsf{Asym}(R_1), \\
\mathsf{Ref}(R_1), & \mathsf{Irref}(R_1), & \mathsf{Dis}(R_1, R_2), & \mathsf{Tra}(R_1), & \mathsf{Fun}(R_1), \\
R_1 \equiv C_1 \times C_2, & P_1 \equiv P_2, & P_1 \sqsubseteq P_2, & \mathsf{Dis}(P_1, P_2), & \mathsf{Fun}(P_1),
\end{array}
$$

where R_1, R_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role terms, C_1, C_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract concept terms, and P_1, P_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concrete role terms. Any expression of the type $w \sqsubseteq R$, where w is a finite string of $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role terms and R is an $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role term is called a *role inclusion axiom (RIA)*.

Next, a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*TBox* is a set of statements of the types:

$$
\begin{array}{l}
C_1 \equiv C_2,\ C_1 \sqsubseteq C_2,\ C_1 \sqsubseteq \forall R_1.C_2,\ \exists R_1.C_1 \sqsubseteq C_2,\ \geq_n R_1.C_1 \sqsubseteq C_2,\ C_1 \sqsubseteq \leq_n R_1.C_2, \\
t_1 \equiv t_2, \quad t_1 \sqsubseteq t_2, \quad C_1 \sqsubseteq \forall P_1.t_1, \quad \exists P_1.t_1 \sqsubseteq C_1, \quad \geq_n P_1.t_1 \sqsubseteq C_1, \quad C_1 \sqsubseteq \leq_n P_1.t_1,
\end{array}
$$

where C_1, C_2 are $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concept terms, t_1, t_2 datatype terms, R_1 a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role term, and P_1 a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concrete role term. Any statement $C \sqsubseteq D$, with C and D $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concept terms, is a *general concept inclusion axiom (GCI)*.

Finally, a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*ABox* is a set of *individual assertions* of the forms:

$$
a : C_1, \quad (a, b) : R_1, \quad a = b, \quad e_d : t_1, \quad (a, e_d) : P_1,
$$

with a, b individual names, C_1 a $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-concept term, R_1 a $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-abstract role term, d a datatype, e_d a constant in $N_C(d)$, t_1 a datatype term, and P_1 a $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-concrete role term.

The semantics of $\mathcal{DL}^{4,\times}_{\mathbf{D}}$ is based on interpretations $\mathbf{I} = (\Delta^{\mathbf{I}}, \Delta_{\mathbf{D}}, \cdot^{\mathbf{I}})$, where $\Delta^{\mathbf{I}}$ and $\Delta_{\mathbf{D}}$ are non-empty disjoint domains such that $d^{\mathbf{D}} \subseteq \Delta_{\mathbf{D}}$, for every $d \in N_D$, and $\cdot^{\mathbf{I}}$ is an interpretation function. The interpretation of concepts and roles, axioms and assertions is illustrated in Table 1.

Name	Syntax	Semantics
concept	A	$A^{\mathbf{I}} \subseteq \Delta^{\mathbf{I}}$
ab. (resp., cn.) rl.	R (resp., P)	$R^{\mathbf{I}} \subseteq \Delta^{\mathbf{I}} \times \Delta^{\mathbf{I}}$ (resp., $P^{\mathbf{I}} \subseteq \Delta^{\mathbf{I}} \times \Delta_{\mathbf{D}}$)
individual	a	$a^{\mathbf{I}} \in \Delta^{\mathbf{I}}$
nominal	$\{a\}$	$\{a\}^{\mathbf{I}} = \{a^{\mathbf{I}}\}$
dtype (resp., ng.)	d (resp., $\neg d$)	$d^{\mathbf{D}} \subseteq \Delta_{\mathbf{D}}$ (resp., $\Delta_{\mathbf{D}} \setminus d^{\mathbf{D}}$)
negative datatype term	$\neg t_1$	$(\neg t_1)^{\mathbf{D}} = \Delta_{\mathbf{D}} \setminus t_1^{\mathbf{D}}$
datatype terms intersection	$t_1 \sqcap t_2$	$(t_1 \sqcap t_2)^{\mathbf{D}} = t_1^{\mathbf{D}} \cap t_2^{\mathbf{D}}$
datatype terms union	$t_1 \sqcup t_2$	$(t_1 \sqcup t_2)^{\mathbf{D}} = t_1^{\mathbf{D}} \cup t_2^{\mathbf{D}}$
constant in $N_C(d)$	e_d	$e_d^{\mathbf{D}} \in d^{\mathbf{D}}$
data range	$\{e_{d_1}, \ldots, e_{d_n}\}$	$\{e_{d_1}, \ldots, e_{d_n}\}^{\mathbf{D}} = \{e_{d_1}^{\mathbf{D}}\} \cup \ldots \cup \{e_{d_n}^{\mathbf{D}}\}$
data range	ψ_d	$\psi_d^{\mathbf{D}}$
data range	$\neg dr$	$\Delta_{\mathbf{D}} \setminus dr^{\mathbf{D}}$
top (resp., bot.)	$\top$ (resp., $\bot$)	$\Delta^{\mathbf{I}}$ (resp., $\emptyset$)
negation	$\neg C$	$(\neg C)^{\mathbf{I}} = \Delta^{\mathbf{I}} \setminus C$
conj. (resp., disj.)	$C \sqcap D$ (resp., $C \sqcup D$)	$(C \sqcap D)^{\mathbf{I}} = C^{\mathbf{I}} \cap D^{\mathbf{I}}$ (resp., $(C \sqcup D)^{\mathbf{I}} = C^{\mathbf{I}} \cup D^{\mathbf{I}}$)
valued exist. quantification	$\exists R.a$	$(\exists R.a)^{\mathbf{I}} = \{x \in \Delta^{\mathbf{I}} : \langle x, a^{\mathbf{I}} \rangle \in R^{\mathbf{I}}\}$
datatyped exist. quantif.	$\exists P.e_d$	$(\exists P.e_d)^{\mathbf{I}} = \{x \in \Delta^{\mathbf{I}} : \langle x, e_d^{\mathbf{D}} \rangle \in P^{\mathbf{I}}\}$
self concept	$\exists R.Self$	$(\exists R.Self)^{\mathbf{I}} = \{x \in \Delta^{\mathbf{I}} : \langle x, x \rangle \in R^{\mathbf{I}}\}$
nominals	$\{a_1, \ldots, a_n\}$	$\{a_1, \ldots, a_n\}^{\mathbf{I}} = \{a_1^{\mathbf{I}}\} \cup \ldots \cup \{a_n^{\mathbf{I}}\}$
universal role	U	$(U)^{\mathbf{I}} = \Delta^{\mathbf{I}} \times \Delta^{\mathbf{I}}$
inverse role	R^-	$(R^-)^{\mathbf{I}} = \{\langle y, x \rangle \mid \langle x, y \rangle \in R^{\mathbf{I}}\}$
concept cart. prod.	$C_1 \times C_2$	$(C_1 \times C_2)^I = C_1^I \times C_2^I$
abstract role complement	$\neg R$	$(\neg R)^{\mathbf{I}} = (\Delta^{\mathbf{I}} \times \Delta^{\mathbf{I}}) \setminus R^{\mathbf{I}}$
abstract role union	$R_1 \sqcup R_2$	$(R_1 \sqcup R_2)^{\mathbf{I}} = R_1^{\mathbf{I}} \cup R_2^{\mathbf{I}}$
abstract role intersection	$R_1 \sqcap R_2$	$(R_1 \sqcap R_2)^{\mathbf{I}} = R_1^{\mathbf{I}} \cap R_2^{\mathbf{I}}$
abstract role domain restr.	$R_{C\|}$	$(R_{C\|})^{\mathbf{I}} = \{\langle x, y \rangle \in R^{\mathbf{I}} : x \in C^{\mathbf{I}}\}$
concrete role complement	$\neg P$	$(\neg P)^{\mathbf{I}} = (\Delta^{\mathbf{I}} \times \Delta^{\mathbf{D}}) \setminus P^{\mathbf{I}}$

concrete role union	$P_1 \sqcup P_2$	$(P_1 \sqcup P_2)^{\mathbf{I}} = P_1^{\mathbf{I}} \cup P_2^{\mathbf{I}}$
concrete role intersection	$P_1 \sqcap P_2$	$(P_1 \sqcap P_2)^{\mathbf{I}} = P_1^{\mathbf{I}} \cap P_2^{\mathbf{I}}$
concrete role domain restr.	$P_{C\mid}$	$(P_{C\mid})^{\mathbf{I}} = \{\langle x,y\rangle \in P^{\mathbf{I}} : x \in C^{\mathbf{I}}\}$
concrete role range restr.	$P_{\mid t}$	$(P_{\mid t})^{\mathbf{I}} = \{\langle x,y\rangle \in P^{\mathbf{I}} : y \in t^{\mathbf{D}}\}$
concrete role restriction	$P_{C_1\mid t}$	$(P_{C_1\mid t})^{\mathbf{I}} = \{\langle x,y\rangle \in P^{\mathbf{I}} : x \in C_1^{\mathbf{I}} \wedge y \in t^{\mathbf{D}}\}$
concept subsum.	$C_1 \sqsubseteq C_2$	$\mathbf{I} \models_{\mathbf{D}} C_1 \sqsubseteq C_2 \iff C_1^{\mathbf{I}} \subseteq C_2^{\mathbf{I}}$
ab. role subsum.	$R_1 \sqsubseteq R_2$	$\mathbf{I} \models_{\mathbf{D}} R_1 \sqsubseteq R_2 \iff R_1^{\mathbf{I}} \subseteq R_2^{\mathbf{I}}$
role incl. axiom	$R_1 \ldots R_n \sqsubseteq R$	$\mathbf{I} \models_{\mathbf{D}} R_1 \ldots R_n \sqsubseteq R \iff R_1^{\mathbf{I}} \circ \ldots \circ R_n^{\mathbf{I}} \subseteq R^{\mathbf{I}}$
cn. role subsum.	$P_1 \sqsubseteq P_2$	$\mathbf{I} \models_{\mathbf{D}} P_1 \sqsubseteq P_2 \iff P_1^{\mathbf{I}} \subseteq P_2^{\mathbf{I}}$
symmetric role	$\mathsf{Sym}(R)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Sym}(R) \iff (R^-)^{\mathbf{I}} \subseteq R^{\mathbf{I}}$
asymmetric role	$\mathsf{Asym}(R)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Asym}(R) \iff R^{\mathbf{I}} \cap (R^-)^{\mathbf{I}} = \emptyset$
transitive role	$\mathsf{Tra}(R)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Tra}(R) \iff R^{\mathbf{I}} \circ R^{\mathbf{I}} \subseteq R^{\mathbf{I}}$
disj. ab. role	$\mathsf{Dis}(R_1, R_2)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Dis}(R_1, R_2) \iff R_1^{\mathbf{I}} \cap R_2^{\mathbf{I}} = \emptyset$
reflexive role	$\mathsf{Ref}(R)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Ref}(R) \iff \{\langle x,x\rangle \mid x \in \Delta^{\mathbf{I}}\} \subseteq R^{\mathbf{I}}$
irreflexive role	$\mathsf{Irref}(R)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Irref}(R) \iff R^{\mathbf{I}} \cap \{\langle x,x\rangle \mid x \in \Delta^{\mathbf{I}}\} = \emptyset$
func. ab. role	$\mathsf{Fun}(R)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Fun}(R) \iff (R^-)^{\mathbf{I}} \circ R^{\mathbf{I}} \subseteq \{\langle x,x\rangle \mid x \in \Delta^{\mathbf{I}}\}$
disj. cn. role	$\mathsf{Dis}(P_1, P_2)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Dis}(P_1, P_2) \iff P_1^{\mathbf{I}} \cap P_2^{\mathbf{I}} = \emptyset$
func. cn. role	$\mathsf{Fun}(P)$	$\mathbf{I} \models_{\mathbf{D}} \mathsf{Fun}(p) \iff \langle x,y\rangle \in P^{\mathbf{I}}$ and $\langle x,z\rangle \in P^{\mathbf{I}}$ imply $y = z$
datatype terms equivalence	$t_1 \equiv t_2$	$\mathbf{I} \models_{\mathbf{D}} t_1 \equiv t_2 \iff t_1^{\mathbf{D}} = t_2^{\mathbf{D}}$
datatype terms diseq.	$t_1 \not\equiv t_2$	$\mathbf{I} \models_{\mathbf{D}} t_1 \not\equiv t_2 \iff t_1^{\mathbf{D}} \neq t_2^{\mathbf{D}}$
datatype terms subsum.	$t_1 \sqsubseteq t_2$	$\mathbf{I} \models_{\mathbf{D}} (t_1 \sqsubseteq t_2) \iff t_1^{\mathbf{D}} \subseteq t_2^{\mathbf{D}}$
concept assertion	$a : C_1$	$\mathbf{I} \models_{\mathbf{D}} a : C_1 \iff (a^{\mathbf{I}} \in C_1^{\mathbf{I}})$
agreement	$a = b$	$\mathbf{I} \models_{\mathbf{D}} a = b \iff a^{\mathbf{I}} = b^{\mathbf{I}}$
disagreement	$a \neq b$	$\mathbf{I} \models_{\mathbf{D}} a \neq b \iff \neg(a^{\mathbf{I}} = b^{\mathbf{I}})$
ab. role asser.	$(a,b) : R$	$\mathbf{I} \models_{\mathbf{D}} (a,b) : R \iff \langle a^{\mathbf{I}}, b^{\mathbf{I}}\rangle \in R^{\mathbf{I}}$
cn. role asser.	$(a,e_d) : P$	$\mathbf{I} \models_{\mathbf{D}} (a,e_d) : P \iff \langle a^{\mathbf{I}}, e_d^{\mathbf{D}}\rangle \in P^{\mathbf{I}}$

Table 1: Semantics of $\mathcal{DL}_{\mathbf{D}}^{4,\times}$.

Legenda. ab: abstract, cn.: concrete, rl.: role, ind.: individual, d. cs.: datatype constant, dtype: datatype, ng.: negated, bot.: bottom, incl.: inclusion, asser.: assertion.

Let $\mathcal{KB} = (\mathcal{R}, \mathcal{T}, \mathcal{A})$ be a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base. An interpretation $\mathbf{I} = (\Delta^{\mathbf{I}}, \Delta_{\mathbf{D}}, \cdot^{\mathbf{I}})$ is a $\mathbf{D}$-model of $\mathcal{R}$ (and write $\mathbf{I} \models_{\mathbf{D}} \mathcal{R}$) if $\mathbf{I}$ satisfies each axiom in $\mathcal{R}$ according to the semantic rules in [5, Table 1]. Similar definitions hold for $\mathcal{T}$ and $\mathcal{A}$ too. Then $\mathbf{I}$ satisfies $\mathcal{KB}$ (and write $\mathbf{I} \models_{\mathbf{D}} \mathcal{KB}$) if it is a $\mathbf{D}$-model of $\mathcal{R}$, $\mathcal{T}$, and $\mathcal{A}$. A knowledge base is *consistent* if it is satisfied by some interpretation.

3 Conjunctive Query Answering for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$

Let $\mathcal{V} = \{v_1, v_2, \ldots\}$ be a denumerable and infinite set of variables disjoint from **Ind** and from $\bigcup\{N_C(d) : d \in N_{\mathbf{D}}\}$. A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*atomic formula* is an expression of of the following types

$$R(w_1, w_2), \quad P(w_1, u_1), \quad C(w_1), \quad w_1 = w_2, \quad u_1 = u_2,$$

where $w_1, w_2 \in \mathcal{V} \cup \mathbf{Ind}$, $u_1, u_2 \in \mathcal{V} \cup \bigcup\{N_C(d) : d \in N_{\mathbf{D}}\}$, R is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-abstract role term, P is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concrete role term, and C is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-concept term. A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-atomic formula containing no variables is said to be *closed*. A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*literal* is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-atomic formula or its negation. A $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-*conjunctive query* is a conjunction of $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-literals. Let $v_1, \ldots, v_n \in \mathcal{V}$ and $o_1, \ldots, o_n \in \mathbf{Ind} \cup \bigcup\{N_C(d) : d \in N_{\mathbf{D}}\}$. A substitution $\sigma \coloneqq \{v_1/o_1, \ldots, v_n/o_n\}$ is a map such that, for every $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-literal L, $L\sigma$ is obtained from L by replacing the occurrences of $v_1, \ldots, v_n$ in L with $o_1, \ldots, o_n$, respectively. Substitutions can be extended to $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive queries in the usual way. Let $Q \coloneqq (L_1 \wedge \ldots \wedge L_m)$ be a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive query, and $\mathcal{KB}$ a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base. A substitution σ involving *exactly* the variables occurring in Q is a *solution for Q w.r.t. $\mathcal{KB}$* if there exists a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation $\mathbf{I}$ such that $\mathbf{I} \models_{\mathbf{D}} \mathcal{KB}$ and $\mathbf{I} \models_{\mathbf{D}} Q\sigma$. The collection Σ of the solutions for Q w.r.t. $\mathcal{KB}$ is the *answer set of Q w.r.t. $\mathcal{KB}$*. Then the *conjuntive query answering* (CQA) problem for Q w.r.t. $\mathcal{KB}$ consists in finding the answer set Σ of Q w.r.t. $\mathcal{KB}$.

We shall solve the CQA problem just stated by reducing it to the analogous problem formulated in the context of the fragment $\mathsf{4LQS^R}$ (and in turn to the decision procedure for $\mathsf{4LQS^R}$ presented in [4]). The CQA problem for $\mathsf{4LQS^R}$-formulae can be stated as follows. Let ϕ be a $\mathsf{4LQS^R}$-formula and let ψ be a conjunction of $\mathsf{4LQS^R}$-quantifier-free atomic formulae of level 0 of the types

$$x = y, \qquad x \in X^1, \qquad \langle x, y\rangle \in X^3$$

or their negations, such that $\mathtt{Var}_0(\psi) \cap \mathtt{Var}_0(\phi) = \emptyset$ and $\mathtt{Var}_1(\psi) \cup \mathtt{Var}_3(\psi) \subseteq \mathtt{Var}_1(\phi) \cup \mathtt{Var}_3(\phi)$. The *CQA problem for ψ w.r.t. ϕ* consists in computing the *answer set of ψ w.r.t. ϕ*, namely the collection Σ' of all the substitutions $\sigma' \coloneqq \{x_1/y_1, \ldots, x_n/y_n\}$ (where $x_1, \ldots, x_n$ are the distinct variables of level 0 in ψ and $\{y_1, \ldots, y_n\} \subseteq \mathtt{Var}_0(\phi)$) such that $\boldsymbol{\mathcal{M}} \models \phi \wedge \psi\sigma'$, for some $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}$. In view of the decidability of the satisfiability problem for $\mathsf{4LQS^R}$-formulae, the CQA problem for $\mathsf{4LQS^R}$-formulae can be solved effectively. Indeed, given two $\mathsf{4LQS^R}$-formulae ϕ and ψ satisfying the above requirements, to compute the answer set of ψ w.r.t. ϕ, for each candidate substitution $\sigma' \coloneqq \{x_1/y_1, \ldots, x_n/y_n\}$ (with $\{x_1, \ldots, x_n\} = \mathtt{Var}_0(\psi)$ and $\{y_1, \ldots, y_n\} \subseteq \mathtt{Var}_0(\phi)$) one has just to test for satisfiability the $\mathsf{4LQS^R}$-formula $\phi \wedge \psi\sigma'$. Since the number of possible candidate substitutions is $|\mathtt{Var}_0(\phi)|^{|\mathtt{Var}_0(\psi)|}$ and the satisfiability test for $\mathsf{4LQS^R}$-formulae can be carried out in an effective manner, the answer set of ψ w.r.t. ϕ can be computed effectively. Summarizing,

Lemma 1. *The CQA problem for* $\mathsf{4LQS^R}$*-formulae can be solved in an effective manner.* □

The following theorem states that also the CQA problem for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ can be solved effectively.

Theorem 1. *Given a* $\mathcal{DL}_{\mathbf{D}}^{4,\times}$*-knowledge base* $\mathcal{KB}$ *and a* $\mathcal{DL}_{\mathbf{D}}^{4,\times}$*-conjunctive query* Q*, the CQA problem for* Q *w.r.t.* $\mathcal{KB}$ *can be solved in an effective manner.*

We first outline the main ideas and then we provide a formal proof of the theorem.

As remarked above, the CQA problem for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ can be solved via an effective reduction to the CQA problem for $\mathsf{4LQS^R}$-formulae, and then exploiting Lemma 1. The reduction is accomplished through a function θ that maps effectively variables in $\mathcal{V}$, individuals in **Ind**, datatype constants in $\bigcup\{N_C(d) : d \in N_{\mathbf{D}}\}$ into variables of sort 0 (of the $\mathsf{4LQS^R}$-language), etc., $\mathcal{DL}_{\mathbf{D}}^{4,\times}$*-TBoxes*, *-RBoxes*, and *-ABoxes*, and $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive queries into $\mathsf{4LQS^R}$-formulae in conjunctive normal form (CNF), which can be used to map effectively CQA problems from the $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-context into the $\mathsf{4LQS^R}$-context. More specifically, given a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base $\mathcal{KB}$ and a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive query Q, using the function θ we can effectively construct the following $\mathsf{4LQS^R}$-formulae in CNF:

$$\phi_{\mathcal{KB}} \coloneqq \bigwedge_{H\in\mathcal{KB}} \theta(H) \wedge \bigwedge_{i=1}^{12} \xi_i, \qquad \psi_Q \coloneqq \theta(Q)\,.^2$$

Then, if we denote by Σ the answer set of Q w.r.t. $\mathcal{KB}$ and by Σ' the answer set of ψ_Q w.r.t. $\phi_{\mathcal{KB}}$, we have that Σ consists of all substitutions σ (involving exactly the variables occurring in Q) such that $\theta(\sigma) \in \Sigma'$. Since, by Lemma 1, Σ' can be computed effectively, then Σ can be computed effectively too.

We are now ready to provide the proof of Theorem 1.

Proof. As preliminary step, observe that the statements of $\mathcal{KB}$ that need to be considered are the following:

- $C_1 \equiv \top$, $C_1 \equiv \neg C_2$, $C_1 \equiv C_2 \sqcup C_3$, $C_1 \equiv \{a\}$, $C_1 \sqsubseteq \forall R_1.C_2$, $\exists R_1.C_1 \sqsubseteq C_2$, $\geq_n R_1.C_1 \sqsubseteq C_2$, $C_1 \sqsubseteq \leq_n R_1.C_2$, $C_1 \sqsubseteq \forall P_1.t_1$, $\exists P_1.t_1 \sqsubseteq C_1$, $\geq_n P_1.t_1 \sqsubseteq C_1$, $C_1 \sqsubseteq \leq_n P_1.t_1$,
- $R_1 \equiv U$, $R_1 \equiv \neg R_2$, $R_1 \equiv R_2 \sqcup R_3$, $R_1 \equiv R_2^-$, $R_1 \equiv id(C_1)$, $R_1 \equiv R_{2_{C_1|}}$, $R_1 \ldots R_n \sqsubseteq R_{n+1}$, $\mathsf{Ref}(R_1)$, $\mathsf{Irref}(R_1)$, $\mathsf{Dis}(R_1, R_2)$, $\mathsf{Fun}(R_1)$, $R_1 \equiv C_1 \times C_2$,
- $P_1 \equiv P_2$, $P_1 \equiv \neg P_2$, $P_1 \equiv P_2 \sqcup P_3$, $P_1 \sqsubseteq P_2$, $\mathsf{Fun}(P_1)$, $P_1 \equiv P_{2_{C_1|}}$, $P_1 \equiv P_{2_{C_1|t_1}}$, $P_1 \equiv P_{2_{|t_1}}$,
- $a : C_1$, $(a, b) : R_1$, $(a, b) : \neg R_1$, $a = b$, $a \neq b$,

[2] The definition of the function θ is inspired to that of the function τ introduced in the proof of Theorem 1 in [7]. Specifically, θ differs from τ as (i) it allows quantification only on variables of level 0, (ii) it treats Boolean operations on concrete roles and the product of concepts, and (iii) it constructs $\mathsf{4LQS^R}$-formulae in CNF. In addition, the constraints ξ_1–ξ_{12} are similar to the constraints ψ_1–ψ_{12} introduced in the proof of Theorem 1 in [7]; they are introduced to guarantee that each model of $\phi_{\mathcal{KB}}$ can be transformed into a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation.

– $e_d : t_1$, $(a, e_d) : P_1$, $(a, e_d) : \neg P_1$.

We solve the problem of CQA for $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ via a reduction to the problem of CQA for $\mathsf{4LQS^R}$, exploiting the decidability result proved in Lemma 1.

We define a function θ that maps the $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base $\mathcal{KB}$ and the $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive query Q in the $\mathsf{4LQS^R}$-formulae in Conjunctive Normal Form (CNF) $\phi_{\mathcal{KB}}$ and ψ_Q, respectively, and the answer set Σ for Q w.r.t. $\mathcal{KB}$ in a set Σ' of (0 level) substitutions in the $\mathsf{4LQS^R}$ formalism.

We will show that, Σ is the answer set for Q w.r.t. $\mathcal{KB}$ iff Σ is equal to $\Sigma' = \bigcup_{\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}} \Sigma'_{\boldsymbol{\mathcal{M}}}$, where $\Sigma'_{\boldsymbol{\mathcal{M}}}$ is the collection of substitutions σ' such that $\boldsymbol{\mathcal{M}} \models \psi_Q \sigma'$.

The definition of the mapping θ is inspired to the definition of the mapping τ introduced in the proof of Theorem 1 in [7]. Specifically, θ differs from τ because it allows quantification only on variables of level 0, it treats Boolean operations on concrete roles and the product of concepts, and it construct $\mathsf{4LQS^R}$-formulae in CNF. To prepare for the definition of θ, we map injectively individuals a, constants $e_d \in N_C(d)$, and variable $y, z, \ldots \in \mathcal{V}$, into level 0 variables x_a, x_{e_d}, x_y, x_z, the constant concepts $\top$ and $\bot$, datatype terms t, and concept terms C into level 1 variables $X^1_\top$, $X^1_\bot$, X^1_t, X^1_C, respectively, and the universal relation on individuals U, abstract role terms R, and concrete role terms P into level 3 variables X^3_U, X^3_R, and X^3_P, respectively.[3]

Then the mapping θ is defined as follows:

$\theta(C_1 \equiv \top) \coloneqq (\forall z)((\neg(z \in X^1_{C_1}) \vee z \in X^1_\top) \wedge (\neg(z \in X^1_\top) \vee z \in X^1_{C_1}))$,
$\theta(C_1 \equiv \neg C_2) \coloneqq (\forall z)((\neg(z \in X^1_{C_1}) \vee \neg(z \in X^1_{C_2})) \wedge (z \in X^1_{C_2} \vee z \in X^1_{C_1}))$,
$\theta(C_1 \equiv C_2 \sqcup C_3) \coloneqq (\forall z)((\neg(z \in X^1_{C_1}) \vee (z \in X^1_{C_2} \vee z \in X^1_{C_3})) \wedge ((\neg(z \in X^1_{C_2}) \vee z \in X^1_{C_1}) \wedge (\neg(z \in X^1_{C_3}) \vee z \in X^1_{C_1}))$,
$\theta(C_1 \equiv \{a\}) \coloneqq (\forall z)(\neg(z \in X^1_{C_1}) \vee z = x_a) \wedge (\neg(z = x_a) \vee z \in X^1_{C_1})$,
$\theta(C_1 \sqsubseteq \forall R_1.C_2) \coloneqq (\forall z_1)(\forall z_2)(\neg(z_1 \in X^1_{C_1}) \vee (\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee z_2 \in X^1_{C_2}))$,
$\theta(\exists R_1.C_1 \sqsubseteq C_2) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee \neg(z_2 \in X^1_{C_1})) \vee z_1 \in X^1_{C_2})$,
$\theta(C_1 \equiv \exists R_1.\{a\}) \coloneqq (\forall z)((\neg(z \in X^1_{C_1}) \vee \langle z, x_a\rangle \in X^3_{R_1}) \wedge (\neg(\langle z, x_a\rangle \in X^3_{R_1}) \vee z \in X^1_{C_1}))$,
$\theta(C_1 \sqsubseteq \leq_n R_1.C_2) \coloneqq (\forall z)(\forall z_1)\ldots(\forall z_{n+1})(\neg(z \in X^1_{C_1}) \vee (\bigwedge_{i=1}^{n+1}(\neg(z_i \in X_{C_2}) \vee \neg(\langle z, z_i\rangle \in X^3_{R_1}) \vee \bigvee_{i<j} z_i = z_j))$,
$\theta(\geq_n R_1.C_1 \sqsubseteq C_2) \coloneqq (\forall z)(\forall z_1)\ldots(\forall z_n)(\bigwedge_{i=1}^{n}((\neg(z_i \in X^1_{C_1}) \vee \neg(\langle z, z_i\rangle \in X^3_{R_1})) \vee \bigvee_{i<j} z_i = z_j) \vee z \in X^1_{C_2})$,
$\theta(C_1 \sqsubseteq \forall P_1.t_1) \coloneqq (\forall z_1)(\forall z_2)(\neg(z_1 \in X^1_{C_1}) \vee (\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee z_2 \in X^1_{t_1}))$,
$\theta(\exists P_1.t_1 \sqsubseteq C_1) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \neg(z_2 \in X^1_{t_1})) \vee z_1 \in X^1_{C_1})$,

[3] The use of level 3 variables to model abstract and concrete role terms is motivated by the fact that their elements, that is ordered pairs $\langle x, y\rangle$, are encoded in Kuratowski's style as $\{\{x\}, \{x, y\}\}$, namely as collections of sets of objects.

$\theta(C_1 \equiv \exists P_1.\{e_d\}) \coloneqq (\forall z)((\neg(z \in X^1_{C_1}) \vee \langle z, x_{e_d}\rangle \in X^3_{P_1}) \wedge (\neg(\langle z, x_{e_d}\rangle \in X^3_{P_1}) \vee z \in X^1_{C_1})),$

$\theta(C_1 \sqsubseteq \leq_n P_1.t_1) \coloneqq (\forall z)(\forall z_1)\ldots(\forall z_{n+1})(\neg(z \in X^1_{C_1}) \vee (\bigwedge_{i=1}^{n+1}(\neg(z_i \in X_{t_1}) \vee \neg(\langle z, z_i\rangle \in X^3_{P_1}) \vee \bigvee_{i<j} z_i = z_j)),$

$\theta(\geq_n P_1.t_1 \sqsubseteq C_1) \coloneqq (\forall z)(\forall z_1)\ldots(\forall z_n)(\bigwedge_{i=1}^{n}((\neg(z_i \in X^1_{t_1}) \vee \neg(\langle z, z_i\rangle \in X^3_{P_1})) \vee \bigvee_{i<j} z_i = z_j) \vee z \in X^1_{C_1}),$

$\theta(R_1 \equiv U) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee \langle z_1, z_2\rangle \in X^3_U) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_U) \vee \langle z_1, z_2\rangle \in X^3_{R_1})),$

$\theta(R_1 \equiv \neg R_2) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee \neg(\langle z_1, z_2\rangle \in X^3_{R_2})) \wedge (\langle z_1, z_2\rangle \in X^3_{R_2} \vee \neg(\langle z_1, z_2\rangle \in X^3_{R_1}))),$

$\theta(R \equiv C_1 \times C_2) \coloneqq (\forall z_1)(\forall z_2)(\neg(\langle z_1, z_2\rangle \in X^3_R) \vee z_1 \in X^1_{C_1}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_R) \vee z_2 \in X^1_{C_2}) \wedge ((\neg(z_1 \in X^1_{C_1}) \vee \neg(z_2 \in X^1_{C_2})) \vee \langle z_1, z_2\rangle \in X^3_R))$

$\theta(R_1 \equiv R_2 \sqcup R_3) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee (\langle z_1, z_2\rangle \in X^3_{R_2} \vee \langle z_1, z_2\rangle \in X^3_{R_3})) \wedge ((\neg(\langle z_1, z_2\rangle \in X^3_{R_2}) \vee \langle z_1, z_2\rangle \in X^3_{R_1}) \wedge ((\neg(\langle z_1, z_2\rangle \in X^3_{R_3}) \vee \langle z_1, z_2\rangle \in X^3_{R_1})))),$

$\theta(R_1 \equiv R_2^-) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee \langle z_2, z_1\rangle \in X^3_{R_2}) \wedge (\neg(\langle z_2, z_1\rangle \in X^3_{R_2}) \vee \langle z_1, z_2\rangle \in X^3_{R_1})),$

$\theta(R_1 \equiv id(C_1)) \coloneqq (\forall z_1)(\forall z_2)(((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee z_1 \in X^1_{C_1}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee z_2 \in X^1_{C_1}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee z_1 = z_2)) \wedge ((\neg(z_1 \in X^1_{C_1}) \vee \neg(z_2 \in X^1_{C_1}) \vee z_1 \neq z_2) \vee \langle z_1, z_2\rangle \in X^3_{R_1})),$

$\theta(R_1 \equiv R_{2_{C_1|}}) \coloneqq (\forall z_1)(\forall z_2)(((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee \langle z_1, z_2\rangle \in X^3_{R_2}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee z_1 \in X^1_{C_1})) \wedge ((\neg(\langle z_1, z_2\rangle \in X^3_{R_2}) \vee \neg(z_1 \in X^1_{C_1})) \vee \langle z_1, z_2\rangle \in X^3_{R_1})),$

$\theta(R_1 \ldots R_n \sqsubseteq R_{n+1}) \coloneqq (\forall z)(\forall z_1)\ldots(\forall z_n)((\neg(\langle z, z_1\rangle \in X^3_{R_1}) \vee \ldots \vee \neg(\langle z_{n-1}, z_n\rangle \in X^3_{R_n})) \vee \langle z, z_n\rangle \in X^3_{R_{n+1}}),$

$\theta(\mathsf{Ref}(R_1)) \coloneqq (\forall z)(\langle z, z\rangle \in X^3_{R_1}),$

$\theta(\mathsf{Irref}(R_1)) \coloneqq (\forall z)(\neg(\langle z, z\rangle \in X^3_{R_1})),$

$\theta(\mathsf{Fun}(R_1)) \coloneqq (\forall z_1)(\forall z_2)(\forall z_3)((\neg(\langle z_1, z_2\rangle \in X^3_{R_1}) \vee \neg(\langle z_1, z_3\rangle \in X^3_{R_1})) \vee z_2 = z_3),$

$\theta(P_1 \equiv P_2) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_2}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{P_2}) \vee \langle z_1, z_2\rangle \in X^3_{P_1})),$

$\theta(P_1 \equiv \neg P_2) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \neg(\langle z_1, z_2\rangle \in X^3_{P_2})) \wedge (\langle z_1, z_2\rangle \in X^3_{P_2} \vee \langle z_1, z_2\rangle \in X^3_{P_1})),$

$\theta(P_1 \sqsubseteq P_2) \coloneqq (\forall z_1)(\forall z_2)(\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_2}),$

$\theta(\mathsf{Fun}(P_1)) \coloneqq (\forall z_1)(\forall z_2)(\forall z_3)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \neg(\langle z_1, z_3\rangle \in X^3_{P_1}) \vee z_2 = z_3),$

$\theta(P_1 \equiv P_{2_{C_1|}}) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_2}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee z_1 \in X^1_{C_1}) \wedge ((\neg\langle z_1, z_2\rangle \in X^3_{P_2}) \vee \neg(z_1 \in X^1_{C_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_1})),$

$\theta(P_1 \equiv P_{2_{|t_1}}) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_2}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee z_2 \in X^1_{t_1}) \wedge ((\neg(\langle z_1, z_2\rangle \in X^3_{P_2}) \vee \neg(z_2 \in X^1_{t_1})) \vee \langle z_1, z_2\rangle \in X^3_{P_1})),$

$\theta(P_1 \equiv P_{2_{C_1|t_1}}) \coloneqq (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_2}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee z_1 \in X^1_{C_1}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{P_1}) \vee z_2 \in X^1_{t_1}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_{P_2}) \vee \neg(z_1 \in X^1_{C_1}) \vee \neg(z_2 \in X^1_{t_1}) \vee \langle z_1, z_2\rangle \in X^3_{P_1})),$

$\theta(t_1 \equiv t_2) \coloneqq (\forall z)((\neg(z \in X^1_{t_1}) \vee z \in X^1_{t_2}) \wedge (\neg(z \in X^1_{t_2}) \vee z \in X^1_{t_1}))$, $\theta(t_1 \equiv \neg t_2) \coloneqq (\forall z)((\neg(z \in X^1_{t_1}) \vee \neg(z \in X^1_{t_2})) \wedge (z \in X^1_{t_2} \vee z \in X^1_{t_1}))$,
$\theta(t_1 \equiv t_2 \sqcup t_3) \coloneqq (\forall z)((\neg(z \in X^1_{t_1}) \vee (z \in X^1_{t_2} \vee z \in X^1_{t_3})) \wedge ((\neg(z \in X^1_{t_2}) \vee z \in X^1_{t_1}) \wedge (\neg(z \in X^1_{t_3}) \vee z \in X^1_{t_1})))$,
$\theta(t_1 \equiv t_2 \sqcap t_3) \coloneqq (\forall z)((\neg(z \in X^1_{t_1}) \vee (z \in X^1_{t_2} \wedge z \in X^1_{t_3})) \wedge (((\neg(z \in X^1_{t_2}) \vee \neg(z \in X^1_{t_3})) \vee z \in X^1_{t_1}))$,
$\theta(t_1 \equiv \{e_d\}) \coloneqq (\forall z)((\neg(z \in X^1_{t_1}) \vee z = x_{e_d}) \wedge (\neg(z = x_{e_d}) \vee z \in X^1_{t_1}))$,
$\theta(a : C_1) \coloneqq x_a \in X^1_{C_1}$,
$\theta((a,b) : R_1) \coloneqq \langle x_a, x_b \rangle \in X^3_{R_1}$,
$\theta((a,b) : \neg R_1) \coloneqq \neg(\langle x_a, x_b \rangle \in X^3_{R_1})$,
$\theta(a = b) \coloneqq x_a = x_b$, $\theta(a \neq b) \coloneqq \neg(x_a = x_b)$,
$\theta(e_d : t_1) \coloneqq x_{e_d} \in X^1_{t_1}$,
$\theta((a,e_d) : P_1) \coloneqq \langle x_a, x_{e_d} \rangle \in X^3_{P_1}$, $\theta((a,e_d) : \neg P_1) \coloneqq \neg(\langle x_a, x_{e_d} \rangle \in X^3_{P_1})$,
$\theta(\alpha \wedge \beta) \coloneqq \theta(\alpha) \wedge \theta(\beta)$.

The mapping θ for $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-conjuctive queries is defined as follows.

$\theta(R_1(w_1, w_2)) \coloneqq \langle x_{w_1}, x_{w_2} \rangle \in X^3_{R_1}$,
$\theta(P_1(w_1, u_1)) \coloneqq \langle x_{w_1}, x_{u_1} \rangle \in X^3_{P_1}$,
$\theta(C_1(w_1) \coloneqq x_{w_1} \in X^1_{C_1}$,
$\theta(w_1 = w_2) \coloneqq x_{w_1} = x_{w_2}$,
$\theta(u_1 = u_2) \coloneqq x_{u_1} = x_{u_2}$.

To complete, we extend the mapping θ on substitutions $\sigma \coloneqq \{x_1/o_1, \ldots, x_n/o_n\}$, where $x_1, \ldots x_n \in \mathcal{V}$ and $o_1, \ldots, o_n \in \mathbf{Ind} \cup \bigcup\{N_C(d) : d \in N_{\mathbf{D}}\}$.

We put $\theta(\sigma)= \theta(\{x_1/o_1, \ldots, x_n/o_n\}) = \{x_{x_1}/x_{o_1}, \ldots, x_{x_n}/x_{o_n}\} = \sigma'$, where $x_{x_1}, \ldots, x_n$, $x_{o_1}, \ldots, x_{o_n}$ are variables of level 0 in $\mathsf{4LQS^R}$.

Let $\mathcal{KB}$ be our $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-knowledge base, and let $\mathsf{cpt}_{\mathcal{KB}}$, $\mathsf{arl}_{\mathcal{KB}}$, $\mathsf{crl}_{\mathcal{KB}}$, and $\mathsf{ind}_{\mathcal{KB}}$ be, respectively, the sets of concept, of abstract role, of concrete role, and of individual names in $\mathcal{KB}$. Moreover, let $N^{\mathcal{KB}}_D \subseteq N_D$ be the set of datatypes in $\mathcal{KB}$, $N^{\mathcal{KB}}_F$ a restriction of N_F assigning to every $d \in N^{\mathcal{KB}}_{\mathbf{D}}$ the set $N^{\mathcal{KB}}_F(d)$ of facets in $N_F(d)$ and in $\mathcal{KB}$. Analogously, let $N^{\mathcal{KB}}_C$ be a restriction of the function N_C associating to every $d \in N^{\mathcal{KB}}_{\mathbf{D}}$ the set $N^{\mathcal{KB}}_C(d)$ of constants contained in $N_C(d)$ and in $\mathcal{KB}$. Finally, for every datatype $d \in N^{\mathcal{KB}}_D$, let $\mathsf{bf}^{\mathbf{D}}_{\mathcal{KB}}(d)$ be the set of facet expressions for d occurring in $\mathcal{KB}$ and not in $N_F(d) \cup \{\top^d, \bot_d\}$. We assume without loss of generality that the facet expressions in $\mathsf{bf}^{\mathbf{D}}_{\mathcal{KB}}(d)$ are in Conjunctive Normal Form. We define the $\mathsf{4LQS^R}$-formula $\phi_{\mathcal{KB}}$ expressing the consistency of $\mathcal{KB}$ as follows:

$$\phi_{\mathcal{KB}} \coloneqq \bigwedge_{H \in \mathcal{KB}} \theta(H) \wedge \bigwedge_{i=1}^{12} \xi_i \,,$$

where
$\xi_1 \coloneqq (\forall z)((\neg(z \in X^1_{\mathbf{I}}) \vee \neg(z \in X^1_{\mathbf{D}})) \wedge (z \in X^1_{\mathbf{D}} \vee z \in X^1_{\mathbf{I}})) \wedge (\forall z)(z \in X^1_{\mathbf{I}} \vee z \in X^1_{\mathbf{D}}) \wedge \neg(\forall z)\neg(z \in X^1_{\mathbf{I}}) \wedge \neg(\forall z)\neg(z \in X^1_{\mathbf{D}})$,

$$\xi_2 \coloneqq ((\forall z)((\neg(z \in X^1_{\mathbf{I}}) \vee z \in X^1_{\top}) \wedge (\neg(z \in X^1_{\top}) \vee z \in X^1_{\mathbf{I}})) \wedge (\forall z)\neg(z \in X_{\perp}),$$

$$\xi_3 \coloneqq \bigwedge_{A \in \mathsf{cpt}_{\mathcal{KB}}} (\forall z)(\neg(z \in X^1_A) \vee z \in X^1_{\mathbf{I}}),$$

$$\xi_4 \coloneqq (\bigwedge_{d \in N_D^{\mathcal{KB}}} ((\forall z)(\neg(z \in X^1_d) \vee z \in X^1_{\mathbf{D}}) \wedge \neg(\forall z)\neg(z \in X^1_d)) \wedge (\forall z) (\bigwedge_{(d_i, d_j \in N_D^{\mathcal{KB}}, i<j)} ((\neg(z \in X^1_{d_i}) \vee \neg(z \in X^1_{d_j})) \wedge (z \in X^1_{d_j} \vee z \in X^1_{d_i})))),$$

$$\xi_5 \coloneqq \bigwedge_{d \in N_D^{\mathcal{KB}}} ((\forall z)((\neg(z \in X^1_d) \vee z \in X^1_{\top_d}) \wedge (\neg(z \in X^1_{\top_d}) \vee z \in X^1_d) \wedge (\forall z)\neg(z \in X^1_{\perp_d})),$$

$$\xi_6 \coloneqq \bigwedge_{\substack{f_d \in N_F^{\mathcal{KB}}(d),\\ d \in N_D^{\mathcal{KB}}}} (\forall z)(\neg(z \in X^1_{f_d}) \vee z \in X^1_d),$$

$$\xi_7 \coloneqq (\forall z_1)(\forall z_2)((\neg(z_1 \in X^1_{\mathbf{I}}) \vee \neg(z_2 \in X^1_{\mathbf{I}}) \vee \langle z_1, z_2\rangle \in X^3_U) \wedge ((\neg(\langle z_1, z_2\rangle \in X^3_U) \vee z_1 \in X^1_{\mathbf{I}}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_U) \vee z_2 \in X^1_{\mathbf{I}}))),$$

$$\xi_8 \coloneqq \bigwedge_{R \in \mathsf{arl}_{\mathcal{KB}}} (\forall z_1)(\forall z_2)((\neg(\langle z_1, z_2\rangle \in X^3_R) \vee z_1 \in X^1_{\mathbf{I}}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_R) \vee z_2 \in X^1_{\mathbf{I}}))),$$

$$\xi_9 \coloneqq \bigwedge_{T \in \mathsf{crl}_{\mathcal{KB}}} (\forall z_1)(\forall z_2)(\neg(\langle z_1, z_2\rangle \in X^3_T) \vee z_1 \in X^1_{\mathbf{I}}) \wedge (\neg(\langle z_1, z_2\rangle \in X^3_T) \vee z_2 \in X^1_{\mathbf{D}}))),$$

$$\xi_{10} \coloneqq \bigwedge_{a \in \mathsf{ind}_{\mathcal{KB}}} (x_a \in X^1_{\mathbf{I}}) \wedge \bigwedge_{\substack{d \in N_D^{\mathcal{KB}},\\ e_d \in N_C^{\mathcal{KB}}(d)}} x_{e_d} \in X^1_d,$$

$$\xi_{11} \coloneqq \bigwedge_{\{e_{d_1}, \ldots, e_{d_n}\} \text{ in } \mathcal{KB}} (\forall z)((\neg(z \in X^1_{\{e_{d_1}, \ldots, e_{d_n}\}}) \vee \bigvee_{i=1}^{n} (z = x_{e_{d_i}})) \wedge (\bigwedge_{i=1}^{n} (z \neq x_{e_{d_i}} \vee z \in X^1_{\{e_{d_1}, \ldots, e_{d_n}\}}))) \wedge \bigwedge_{\{a_1, \ldots, a_n\} \text{ in } \mathcal{KB}} (\forall z)((\neg(z \in X^1_{\{a_1, \ldots, a_n\}}) \vee \bigvee_{i=1}^{n} (z = x_{a_i})) \wedge (\bigwedge_{i=1}^{n} (z \neq x_{a_i} \vee z \in X^1_{\{a_1, \ldots, a_n\}}))),$$

$$\xi_{12} \coloneqq \bigwedge_{\substack{d \in N_{\mathbf{D}}^{\mathcal{KB}},\\ \psi_d \in \mathsf{bf}^{\mathbf{D}}_{\mathcal{KB}}(d)}} (\forall z)(\neg(z \in X^1_{\psi_d}) \vee z \in \zeta(X^1_{\psi_d})) \wedge (\neg(z \in \zeta(X^1_{\psi_d})) \vee z \in X^1_{\psi_d})$$

with ζ the transformation function from $\mathsf{4LQS^R}$-variables of level 1 to $\mathsf{4LQS^R}$-formulae recursively defined, for $d \in N_{\mathbf{D}}^{\mathcal{KB}}$, by

$$\zeta(X^1_{\psi_d}) \coloneqq \begin{cases} X^1_{\psi_d} & \text{if } \psi_d \in N_F^{\mathcal{KB}}(d) \cup \{\top^d, \bot_d\} \\ \neg\zeta(X^1_{\chi_d}) & \text{if } \psi_d = \neg\chi_d \\ \zeta(X^1_{\chi_d}) \wedge \zeta(X^1_{\varphi_d}) & \text{if } \psi_d = \chi_d \wedge \varphi_d \\ \zeta(X^1_{\chi_d}) \vee \zeta(X^1_{\varphi_d}) & \text{if } \psi_d = \chi_d \vee \varphi_d\,. \end{cases}$$

In the above formulae, the variable $X^1_{\mathbf{I}}$ denotes the set of individuals $\mathbf{Ind}$, X^1_d a datatype $d \in N_D^{\mathcal{KB}}$, $X^1_{\mathbf{D}}$ a superset of the union of datatypes in $N_D^{\mathcal{KB}}$, $X^1_{\top_d}$ and $X^1_{\bot_d}$ the constants $\top_d$ and $\bot_d$, and $X^1_{f_d}$, $X^1_{\psi_d}$ a facet f_d and a facet expression ψ_d, for $d \in N_D^{\mathcal{KB}}$, respectively. In addition, X^1_A, X^3_R, X^3_T denote a concept name A, an abstract role name R, and a concrete role name T occurring in $\mathcal{KB}$, respectively. Finally, $X^1_{\{e_{d_1},\ldots,e_{d_n}\}}$ denotes a data range $\{e_{d_1},\ldots,e_{d_n}\}$ occurring in $\mathcal{KB}$, and $X^1_{\{a_1,\ldots,a_n\}}$ a finite set $\{a_1,\ldots,a_n\}$ of nominals in $\mathcal{KB}$.

The constraints $\xi_1-\xi_{12}$, slightly different from the constraints $\psi_1-\psi_{12}$ defined in the proof of Theorem 1 in [7], are introduced to guarantee that each model of $\phi_{\mathcal{KB}}$ can be easily transformed in a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation. To prove the theorem, we show that Σ is the answer set for Q w.r.t. $\mathcal{KB}$ iff Σ is equal to $\bigcup_{\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}} \Sigma'_{\boldsymbol{\mathcal{M}}}$, where $\Sigma'_{\boldsymbol{\mathcal{M}}}$ is the collection of substitutions σ such that $\boldsymbol{\mathcal{M}} \models \psi_Q\sigma$.

Preliminarly we show that if $\boldsymbol{\mathcal{M}}$ is a $\mathsf{4LQS^R}$-interpretation such that $\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}$, we can construct a $\mathcal{DL}_{\mathbf{D}}^{4}$-interpretation $\mathbf{I}_{\boldsymbol{\mathcal{M}}}$ such that $\mathbf{I}_{\boldsymbol{\mathcal{M}}} \models_{\mathbf{D}} \mathcal{KB}$ and, if $\mathbf{I}$ is a $\mathcal{DL}_{\mathbf{D}}^{4}$-interpretation such that $\mathbf{I} \models_{\mathbf{D}} \mathcal{KB}$, we can construct a $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}_{\mathbf{I}}$ such that $\boldsymbol{\mathcal{M}}_{\mathbf{I}} \models \phi_{\mathcal{KB}}$. Thus, let $\boldsymbol{\mathcal{M}}$ be any $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}$ such that $\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}$. Reasoning as in [7], it is not hard to see that such $\boldsymbol{\mathcal{M}}$ is a $\mathsf{4LQS^R}$-interpretation of the form $\boldsymbol{\mathcal{M}} = (D_1 \cup D_2, M)$, where

- D_1 and D_2 are disjoint nonempty sets and $\bigcup_{d \in N_D^{\mathcal{K}}} d^{\mathbf{D}} \subseteq D_2$,
- $MX^1_{\mathbf{I}} \coloneqq D_1$, $MX^1_{\mathbf{D}} \coloneqq D_2$, $MX^1_d \coloneqq d^{\mathbf{D}}$, for every $d \in N_D^{\mathcal{K}}$,
- $MX^1_{f_d} \coloneqq f_d^{\mathbf{D}}$, for every $f_d \in N_F^{\mathcal{K}}(d)$, with $d \in N_D^{\mathcal{K}}$.

Exploiting the fact that $\boldsymbol{\mathcal{M}}$ satisfies the constrains $\xi_1-\xi_{12}$, it is then possible to define a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation $\mathbf{I}_{\mathcal{M}} = (\Delta^I, \Delta_{\mathbf{D}}, \cdot^I)$, by putting

- $\Delta^{\mathbf{I}} \coloneqq MX^1_{\mathbf{I}}$,
- $\Delta_{\mathbf{D}} \coloneqq MX^1_{\mathbf{D}}$,
- $A^{\mathbf{I}} \coloneqq MX^1_A$, for every concept name $A \in \mathsf{cpt}_{\mathcal{KB}}$,
- $S^{\mathbf{I}} \coloneqq \{\langle u_1, u_2\rangle : u_1 \in MX^1_I, u_2 \in MX^1_I, \langle u_1, u_2\rangle \in MX^3_S\}$, for every abstract role name $S \in \mathsf{arl}_{\mathcal{KB}}$,
- $T^{\mathbf{I}} \coloneqq \{\langle u_1, u_2\rangle : u_1 \in MX^1_I, u_2 \in MX^1_D, \langle u_1, u_2\rangle \in MX^3_T\}$, for every concrete role name $T \in \mathsf{crl}_{\mathcal{KB}}$,
- $a^{\mathbf{I}} \coloneqq Mx_a$, for every individual $a \in \mathsf{ind}_{\mathcal{KB}}$,
- $e_d^D \coloneqq Mx_{e_d}$, for every constant $e_d \in N_C^{\mathcal{KB}}(d)$ with $d \in N_D^{\mathcal{KB}}$.

Since $\boldsymbol{\mathcal{M}} \models \bigwedge_{H \in \mathcal{KB}} \theta(H) \wedge \bigwedge_{i=1}^{12} \xi_i$, and, as it can easily checked, $\mathbf{I}_{\mathcal{M}} \models_{\mathbf{D}} H$, iff $\boldsymbol{\mathcal{M}} \models \theta(H)$, for every statement $H \in \mathcal{KB}$, we plainly have that $\mathbf{I}_{\mathcal{M}} \models_{\mathbf{D}} \mathcal{KB}$. Conversely, let $\mathbf{I} = (\Delta^{\mathbf{I}}, \Delta_{\mathbf{D}}, \cdot^{\mathbf{I}})$ be a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation such that $\mathbf{I} \models_{\mathbf{D}} \mathcal{KB}$. We show how to construct, out of the datatype map $\mathbf{D}$ and the $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation $\mathbf{I}$, a $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}_{\mathbf{I},\mathbf{D}} = (D_{\mathbf{I},\mathbf{D}}, M_{\mathbf{I},\mathbf{D}})$ which satisfies $\phi_{\mathcal{KB}}$. Let us put $D_{\mathbf{I},\mathbf{D}} := \Delta^{\mathbf{I}} \cup \Delta_{\mathbf{D}}$ and define $M_{\mathbf{I},\mathbf{D}}$ by putting $M_{\mathbf{I},\mathbf{D}} X^1_{\mathbf{I}} := \Delta^{\mathbf{I}}$, $M_{\mathbf{I},\mathbf{D}} X^1_{\mathbf{D}} := \Delta_{\mathbf{D}}$, $M_{\mathbf{I},\mathbf{D}} X^3_U := U^{\mathbf{I}}$, $M_{\mathbf{I},\mathbf{D}} X^1_{dr} := dr^{\mathbf{D}}$, for every variable X^1_{dr} in $\phi_{\mathcal{KB}}$ denoting a data range dr occurring in $\mathcal{KB}$, $M_{\mathbf{I},\mathbf{D}} X^1_A := A^{\mathbf{I}}$, for every X^1_A in $\phi_{\mathcal{KB}}$ denoting a concept name in $\mathcal{KB}$, and $M_{\mathbf{I},\mathbf{D}} X^3_S := S^{\mathbf{I}}$, for every X^3_S in $\phi_{\mathcal{KB}}$ denoting an abstract role name in $\mathcal{KB}$. Variables X^3_T, denoting concrete role names, and variables x_a, x_{e_d}, denoting individuals and datatype constants, respectively, are interpreted in a similar way. From the definitions of $\mathbf{D}$ and $\mathbf{I}$, it follows easily that $\boldsymbol{\mathcal{M}}_{\mathbf{I},\mathbf{D}}$ satisfies the formulae ξ_1-ξ_{12} and $\theta(H)$, for every statement $H \in \mathcal{KB}$, and, therefore, that $\boldsymbol{\mathcal{M}}_{\mathbf{I},\mathbf{D}}$ is a model for $\phi_{\mathcal{KB}}$.

Now we prove the first part of the theorem. Let us assume that Σ is the answer set for Q w.r.t. $\mathcal{KB}$. We have to show that Σ is equal to $\Sigma' = \bigcup_{\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}} \Sigma'_{\boldsymbol{\mathcal{M}}}$, where $\Sigma'_{\boldsymbol{\mathcal{M}}}$ is the collection of all the substitutions σ' such that $\boldsymbol{\mathcal{M}} \models \psi_Q \sigma'$.

By contradiction, let us assume that there exists a $\sigma \in \Sigma$ such that $\sigma \notin \Sigma'$, namely $\boldsymbol{\mathcal{M}} \not\models \psi_Q \sigma$, for every $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}$ with $\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}$. Since $\sigma \in \Sigma$ there is a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretation $\mathbf{I}$ such that $\mathbf{I} \models_{\mathbf{D}} \mathcal{KB}$ and $\mathbf{I} \models_{\mathbf{D}} Q\sigma$. Then, by the construction above, we can define a $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}_{\mathbf{I}}$ such that $\boldsymbol{\mathcal{M}}_{\mathbf{I}} \models \phi_{\mathcal{KB}}$ and $\boldsymbol{\mathcal{M}}_{\mathbf{I}} \models \psi_Q \theta \sigma$. Absurd.

Conversely, let $\sigma' \in \Sigma'$ and assume by contradiction that $\sigma' \notin \Sigma$. Then, for all $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-interpretations such that $\mathbf{I} \models_D \mathcal{KB}$, it holds that $\mathbf{I} \not\models_D Q\sigma'$. Since $\sigma' \in \Sigma'$, there is a $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}$ such that $\boldsymbol{\mathcal{M}} \models \phi_{\mathcal{KB}}$ and $\boldsymbol{\mathcal{M}} \models \psi\sigma'$. Then, by the construction above, we can define a $\mathcal{DL}_{\mathbf{D}}^{4}$-interpretation $\mathbf{I}_{\boldsymbol{\mathcal{M}}}$ such that $\mathbf{I}_{\boldsymbol{\mathcal{M}}} \models_D \mathcal{KB}$ and $\mathbf{I}_{\boldsymbol{\mathcal{M}}} \models_D Q\sigma'$. Absurd.

4 A tableau-based procedure

In this section, we illustrate a KE-tableau based procedure that, given a $\mathsf{4LQS^R}$-formula $\phi_{\mathcal{KB}}$ corresponding to a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base and a $\mathsf{4LQS^R}$-formula ψ_Q corresponding to a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive query Q, yields all the substitutions $\sigma = \{x_1/y_1, \ldots, x_n/y_n\}$, with $\{x_1, \ldots, x_n\} = \mathtt{Var}_0(\psi_Q)$ and $\{y_1, \ldots, y_n\} \subseteq \mathtt{Var}_0(\phi_{\mathcal{KB}})$, belonging to the answer set Σ' of ψ_Q w.r.t. $\phi_{\mathcal{KB}}$.

Let $\overline{\phi}_{\mathcal{KB}}$ be the formula obtained from $\phi_{\mathcal{KB}}$ by:

- moving universal quantifiers in $\phi_{\mathcal{KB}}$ as inwards as possible according to the rule $(\forall z)(A(z) \wedge B(z)) \longleftrightarrow ((\forall z)A(z) \wedge (\forall z)B(z))$,
- renaming universally quantified variables so as to make them pairwise distinct.

Let $F_1, \ldots, F_k$ be the conjuncts of $\overline{\phi}_{\mathcal{KB}}$ that are $\mathsf{4LQS^R}$-quantifier-free atomic formulae and $S_1, \ldots, S_m$ the conjuncts of $\overline{\phi}_{\mathcal{KB}}$ that are $\mathsf{4LQS^R}$-purely universal formulae. For every $S_i = (\forall z^i_1) \ldots (\forall z^i_{n_i}) \chi_i$, $i = 1, \ldots, m$, we put

$$Exp(S_i) \coloneqq \bigwedge_{\{x_{a_1},\ldots,x_{a_{n_i}}\}\subseteq \mathtt{Var}_0(\phi_{\mathcal{KB}})} S_i\{z_1^i/x_{a_1},\ldots,z_{n_i}^i/x_{a_{n_i}}\}.$$

Let $\Phi_{\mathcal{KB}} \coloneqq \{F_j : i = 1,\ldots,k\} \cup \bigcup_{i=1}^{m} Exp(S_i)$.

To prepare for the KE-tableau based procedure to be described next, we introduce some useful notions and notations (see [11] for a detailed overview of KE-tableau, an optimized variant of semantic tableaux).

Let $\Phi = \{C_1,\ldots,C_p\}$ be a collection of disjunctions of $\mathsf{4LQS^R}$-quantifier-free atomic formulae of level 0 of the types: $x = y$, $x \in X^1$, $\langle x,y\rangle \in X^3$. $\mathcal{T}$ is a *KE-tableau* for Φ if there exists a finite sequence $\mathcal{T}_1,\ldots,\mathcal{T}_t$ such that (i) $\mathcal{T}_1$ is a one-branch tree consisting of the sequence $C_1,\ldots,C_p$, (ii) $\mathcal{T}_t = \mathcal{T}$, and (iii) for each $i < t$, $\mathcal{T}_{i+1}$ is obtained from $\mathcal{T}_i$ by an application of one of the rules in Fig 1. The set of formulae $\mathcal{S}_i^{\overline{\beta}} = \{\overline{\beta}_1,\ldots,\overline{\beta}_n\} \setminus \{\overline{\beta}_i\}$ occurring as premise in the E-rule contains the complements of all the components of the formula β with the exception of the component β_i.

$$\frac{\beta_1 \vee \ldots \vee \beta_n \qquad \mathcal{S}_i^{\overline{\beta}}}{\beta_i}\ \textbf{E-Rule} \qquad\qquad \frac{}{A \ \mid\ \overline{A}}\ \textbf{PB-Rule}$$

where $\mathcal{S}_i^{\overline{\beta}} \coloneqq \{\overline{\beta}_1,...,\overline{\beta}_n\} \setminus \{\overline{\beta}_i\}$, for $i = 1,...,n$ — with A a literal

Fig. 1. Expansion rules for the KE-tableau.

Let $\mathcal{T}$ be a KE-tableau. A branch θ of $\mathcal{T}$ is *closed* if it contains both A and $\neg A$, for some formula A. Otherwise, the branch is *open*. A formula $\beta_1 \vee \ldots \vee \beta_n$ is *fulfilled* in a branch ϑ, if β_i is in θ, for some $i = 1,\ldots,n$. A branch ϑ is *complete* if every formula $\beta_1 \vee \ldots \vee \beta_n$ occurring in ϑ is fulfilled. A KE-tableau is *complete* if all its branches are complete.

Next we introduce the procedure Saturate-KB that takes as input the set $\Phi_{\mathcal{KB}}$ constructed from a $\mathsf{4LQS^R}$-formula $\phi_{\mathcal{KB}}$ representing a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base $\mathcal{KB}$ as shown above, and yields a complete KE-tableau $\mathcal{T}_{\mathcal{KB}}$ for $\Phi_{\mathcal{KB}}$.

Procedure 1 *Saturate-KB($\Phi_{\mathcal{KB}}$)*

1. $\mathcal{T}_{\mathcal{KB}} := \Phi_{\mathcal{KB}}$;

2. Select an open branch ϑ of $\mathcal{T}_{\mathcal{KB}}$ that is not yet complete.
   (a) Select a formula $\beta_1 \vee \ldots \vee \beta_n$ on ϑ that is not fulfilled.
   (b) If $\mathcal{S}_j^{\overline{\beta}}$ is in ϑ, for some $j \in \{1,\ldots,n\}$, apply the E-Rule to $\beta_1 \vee \ldots \vee \beta_n$ and $\mathcal{S}_j^{\overline{\beta}}$ on ϑ and go to step 2.
   (c) If $\mathcal{S}_j^{\overline{\beta}}$ is not in ϑ, for every $j = 1,\ldots,n$, let $B^{\overline{\beta}}$ be the collection of formulae $\overline{\beta}_1,\ldots,\overline{\beta}_n$ present in ϑ and let $\overline{\beta}_h$ be the lowest index formula such that $\overline{\beta}_h \in \{\{\overline{\beta}_1,\ldots,\overline{\beta}_n\} \setminus B^{\overline{\beta}}\}$, then apply the PB-rule to $\overline{\beta}_h$ on ϑ, and go to step 2.

3. Return $\mathcal{T}_{\mathcal{KB}}$.

Soundness of Procedure 1 can be easily proved in a standard way and its completeness can be shown much along the lines of Proposition 36 in [11]. Concerning termination of Procedure 1, our proof is based on the following two facts. The rules in Fig. 1 are applied only to non-fulfilled formulae on open branches and tend to reduce the number of non-fulfilled formulae occurring on the considered branch. In particular, when the E-Rule is applied on a branch ϑ, the number of non-fulfilled formulae on ϑ decreases. In case of application of the PB-Rule on a formula $\beta = \beta_1 \vee \ldots \vee \beta_n$ on a branch, the rule generates two branches. In one of them the number of non-fulfilled formulae decreases (because β becomes fulfilled). In the other one the number of non-fulfilled formulae stays constant but the subset $B^{\overline{\beta}}$ of $\{\overline{\beta}_1, \ldots, \overline{\beta}_n\}$ occurring on the branch gains a new element. Once $|B^{\overline{\beta}}|$ gets equal to $n-1$, namely after at most $n-1$ applications of the PB-rule, the E-rule is applied and the formula $\beta = \beta_1 \vee \ldots \vee \beta_n$ becomes fulfilled, thus decrementing the number of non-fulfilled formulae on the branch. Since the number of non-fulfilled formulae on each open branch gets equal to zero after a finite number of steps and the rules of Fig. 1 can be applied only to non-fulfilled formulae on open branches, the procedure terminates.

By the completeness of Procedure 1, each branch ϑ of $\mathcal{T}_{\mathcal{KB}}$ induces a $\mathsf{4LQS^R}$-interpretation $\boldsymbol{\mathcal{M}}_\vartheta$ such that $\boldsymbol{\mathcal{M}}_\vartheta \models \Phi_{\mathcal{KB}}$. We define $\boldsymbol{\mathcal{M}}_\vartheta = (D_\vartheta, M_\vartheta)$ as follows. We put

- $D_\vartheta \coloneqq \{x \in \mathcal{V}_0 : x \text{ occurs in } \vartheta\}$;
- $M_\vartheta x \coloneqq x$, for every $x \in D_\vartheta$;
- $M_\vartheta X^1_C = \{x : x \in X^1_C \text{ is in } \vartheta\}$, for every $X^1_C \in \mathcal{V}_1$ occurring ϑ;
- $M_\vartheta X^3_R = \{\langle x,y\rangle : \langle x,y\rangle \in X^3_R \text{ is in } \vartheta\}$, for every $X^3_R \in \mathcal{V}_3$ occurring in ϑ.

It is easy to check that $\boldsymbol{\mathcal{M}}_\vartheta \models \overline{\phi}_{\mathcal{KB}}$ and thus, plainly, that $\boldsymbol{\mathcal{M}}_\vartheta \models \phi_{\mathcal{KB}}$. Next, we provide some complexity results. Let r be the maximum number of universal quantifiers in S_i, and $k \coloneqq |\mathtt{Var}_0(\overline{\phi}_{\mathcal{KB}})|$. Then, each S_i generates k^r expansions. Since the knowledge base contains m such formulae, the number of disjunctions in the initial branch of the KE-tableau is $m \cdot k^r$. Next, let ℓ be the maximum number of literals in S_i, for $i = 1, \ldots, m$. Then, the maximum depth of the KE-tableau, namely the maximum size of the models of $\Phi_{\mathcal{KB}}$ constructed as illustrated above, is $\mathcal{O}(\ell m k^r)$ and the number of leaves of the tableau, that is the number of such models of $\Phi_{\mathcal{KB}}$, is $O(2^{\ell m k^r})$.

We now describe a procedure that, given a KE-tableau constructed by Procedure 1 and a $\mathsf{4LQS^R}$-formula ψ_Q representing a $\mathcal{DL}^{4,\times}_{\mathbf{D}}$-conjunctive query Q, yields all the substitutions σ' in the answer set Σ' of ψ_Q w.r.t. $\phi_{\mathcal{KB}}$. By the soundness of Procedure 1, we can limit ourselves to consider only the models $\boldsymbol{\mathcal{M}}_\vartheta$ of $\phi_{\mathcal{KB}}$ induced by each open branch ϑ of $\mathcal{T}_{\mathcal{KB}}$. For every open and complete branch ϑ of $\mathcal{T}_{\mathcal{KB}}$, we construct a decision tree $\mathcal{D}_\vartheta$ such that every maximal branch of $\mathcal{D}_\vartheta$ defines a substitution σ' such that $\boldsymbol{\mathcal{M}}_\vartheta \models \psi_Q\sigma'$.

Let d be the number of literals in ψ_Q. $\mathcal{D}_\vartheta$ is a finite labelled tree of depth $d+1$ whose labelling satisfies the following conditions, for $i = 0, \ldots, d$:

(i) every node of $\mathcal{D}_\vartheta$ at level i is labelled with $(\sigma_i, \psi_Q\sigma_i)$, and, in particular, the root is labelled with $(\sigma'_0, \psi_Q\sigma'_0)$, where σ'_0 is the empty substitution;

(ii) if a node at level i is labelled with $(\sigma'_i, \psi_Q\sigma'_i)$, then its s-successors, with $s > 0$, are labelled with

$$\big(\sigma'_i\varrho_1^{q_i+1}, \psi_Q(\sigma'_i\varrho_1^{q_i+1})\big), \ldots, \big(\sigma'_i\varrho_s^{q_i+1}, \psi_Q(\sigma'_i\varrho_s^{q_i+1})\big),$$

where q_{i+1} is the $(i+1)$-st conjunct of $\psi_Q\sigma'_i$ and $\mathcal{S}_{q_{i+1}} = \{\varrho_1^{q_i+1}, \ldots, \varrho_s^{q_i+1}\}$ is the collection of the substitutions $\varrho = \{x_1/y_1, \ldots, x_j/y_j\}$ with $\{x_1, \ldots, x_j\} = \mathtt{Var}_0(q_{i+1})$ such that $p = q_{i+1}\varrho$, for some literal p on ϑ. If $s = 0$, the node labelled with $(\sigma'_i, \psi_Q\sigma'_i)$ is a leaf node and, if $i = d$, σ'_i is added to Σ'.

Let $\delta(\mathcal{T}_{\mathcal{KB}})$ and $\lambda(\mathcal{T}_{\mathcal{KB}})$ be, respectively, the maximum depth of $\mathcal{T}_{\mathcal{KB}}$ and the number of leaves of $\mathcal{T}_{\mathcal{KB}}$ computed above. Plainly, $\delta(\mathcal{T}_{\mathcal{KB}}) = \mathcal{O}(\ell m k^r)$ and $\lambda(\mathcal{T}_{\mathcal{KB}}) = \mathcal{O}(2^{\ell m k^r})$. It is easy to verify that $s = 2^k$ is the maximum branching of $\mathcal{D}_\vartheta$. Since $\mathcal{D}_\vartheta$ is a s-ary tree of depth $d+1$, where d is the number of literals in ψ_Q, and the s-successors of a node are computed in $\mathcal{O}(\delta(\mathcal{T}_{\mathcal{KB}}))$ time, the number of leaves in $\mathcal{D}_\vartheta$ is $\mathcal{O}(s^{(d+1)}) = \mathcal{O}(2^{k(d+1)})$ and they are computed in $\mathcal{O}(2^{k(d+1)}\delta(\mathcal{T}_{\mathcal{KB}}))$ time. Finally, since we have $\lambda(\mathcal{T}_{\mathcal{KB}})$ of such decision trees, the answer set of ψ_Q w.r.t. $\phi_{\mathcal{KB}}$ is computed in time

$$\mathcal{O}(2^{k(d+1)}\delta(\mathcal{T}_{\mathcal{KB}})\lambda(\mathcal{T}_{\mathcal{KB}})) = \mathcal{O}(2^{k(d+1)} \cdot \ell m k^r \cdot 2^{\ell m k^r}) = \mathcal{O}(\ell m k^r 2^{k(d+1)+\ell m k^r}) \ .$$

Since the size of $\phi_{\mathcal{KB}}$ and of ψ_Q are polynomially related to those of $\mathcal{KB}$ and of Q, respectively (see [5] for details on the reduction), the construction of the answer set of Q with respect to $\mathcal{KB}$ can be done in double-exponential time. In case $\mathcal{KB}$ contains no role chain axioms and qualified cardinality restrictions, the complexity of our CQA problem is in EXPTIME, since the maximum number of universal quantifiers in $\phi_{\mathcal{KB}}$, namely r, is a constant (in particular $r = 3$). We remark that such result is comparable with the complexity of the CQA problem for a large family of description logics such as $\mathcal{SHIQ}$ [22]. In particular, the CQA problem for the very expressive description logic $\mathcal{SROIQ}$ turns out to be 2-NEXPTIME-complete.

5 Conclusions

We have introduced the description logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R},\times}\rangle(\mathbf{D})$ ($\mathcal{DL}_{\mathbf{D}}^{4,\times}$, for short) that extends the logic $\mathcal{DL}\langle\mathsf{4LQS}^{\mathsf{R}}\rangle(\mathbf{D})$ with Boolean operations on concrete roles and on the product of concepts. We addressed the problem of Conjunctive Query Answering for the description logic $\mathcal{DL}_{\mathbf{D}}^{4,\times}$ by formalizing $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge bases and $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive queries in terms of formulae of $\mathsf{4LQS}^{\mathsf{R}}$. Such formalization seems to be promising for implementation purposes.

In our approach, we first constructed a KE-tableau $\mathcal{T}_{\mathcal{KB}}$ for $\phi_{\mathcal{KB}}$, a $\mathsf{4LQS}^{\mathsf{R}}$-formalization of a given $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-knowledge base $\mathcal{KB}$, whose branches induce the models of $\phi_{\mathcal{KB}}$. Then we computed the answer set of a $\mathsf{4LQS}^{\mathsf{R}}$-formula ψ_Q, representing a $\mathcal{DL}_{\mathbf{D}}^{4,\times}$-conjunctive query Q, with respect to $\phi_{\mathcal{KB}}$ by means of a forest of decision trees based on the branches of $\mathcal{T}_{\mathcal{KB}}$ and gave some complexity results.

We plan to generalize our procedure with a data-type checker in order to extend reasoning with data-types, and also to extend $\mathsf{4LQS^R}$ with data-type groups. We also intend to improve the efficiency of the knowledge base saturation algorithm and query answering algorithm, and to extend the expressiveness of the queries. Finally, we intend to study a parallel model of the procedure described and to provide an implementation of it.

References

1. D. Calvanese, T. Eiter, and M. Ortiz, "Answering regular path queries in expressive description logics: An automata-theoretic approach," in *AAAI*, vol. 7, 2007, pp. 391–396.
2. D. Calvanese, G. D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati, "Data complexity of query answering in description logics," *Artificial Intelligence*, vol. 195, pp. 335 – 360, 2013.
3. D. Calvanese, G. D. Giacomo, and M. Lenzerini, "On the decidability of query containment under constraints," in *Proc. of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, ser. PODS '98. New York, NY, USA: ACM, 1998, pp. 149–158.
4. D. Cantone and M. Nicolosi-Asmundo, "On the satisfiability problem for a 4-level quantified syllogistic and some applications to modal logic," *Fundamenta Informaticae*, vol. 124, no. 4, pp. 427–448, 2013.
5. D. Cantone, M. Nicolosi-Asmundo, and D. Santamaria, "Conjunctive Query Answering via a Fragment of Set Theory (Extended Version)," 2016, available at https://archive.org/details/CanNicSan16_Extended.
6. D. Cantone and C. Longo, "A decidable two-sorted quantified fragment of set theory with ordered pairs and some undecidable extensions," *Theor. Comput. Sci.*, vol. 560, pp. 307–325, 2014.
7. D. Cantone, C. Longo, M. N. Asmundo, and D. F. Santamaria, "Web ontology representation and reasoning via fragments of set theory," in *Web Reasoning and Rule Systems - 9th International Conference, RR 2015, Berlin, Germany, August 4-5, 2015, Proceedings*, 2015, pp. 61–76.
8. D. Cantone, C. Longo, and M. Nicolosi Asmundo, "A decision procedure for a two-sorted extension of multi-level syllogistic with the Cartesian product and some map constructs," in *Proc. of the 25th Italian Conference on Computational Logic (CILC 2010), Rende, Italy, July 7-9, 2010*, W. Faber and N. Leone, Eds., vol. 598. CEUR Workshop Proceedings, ISSN 1613-0073, June 2010, pp. 1–18 (paper 11).
9. D. Cantone, C. Longo, and M. Nicolosi-Asmundo, "A decidable quantified fragment of set theory involving ordered pairs with applications to description logics," in *Proc. Computer Science Logic, 20th Annual Conf. of the EACSL, CSL 2011, September 12-15, 2011, Bergen, Norway*, 2011, pp. 129–143.
10. D. Cantone, C. Longo, and A. Pisasale, "Comparing description logics with multi-level syllogistics: the description logic $\mathcal{DL}\langle\mathsf{MLSS}^{\times}_{2,m}\rangle\mathcal{DL}\langle\mathsf{MLSS}^{\times}_{2,m}\rangle$," in *6th Workshop on Semantic Web Applications and Perspectives (Bressanone, Italy, Sep. 21-22, 2010)*, P. Traverso, Ed., 2010, pp. 1–13.
11. M. D'Agostino, "Tableau methods for classical propositional logic," in *Handbook of Tableau Methods*, M. D'Agostino, D. M. Gabbay, R. Hähnle, and J. Posegga, Eds. Springer, 1999, pp. 45–123.

12. N. Dershowitz and J.-P. Jouannaud, "Rewrite systems," in *Handbook of Theoretical Computer Science (Vol. B)*, J. van Leeuwen, Ed. Cambridge, MA, USA: MIT Press, 1990, pp. 243–320.
13. B. Glimm, I. Horrocks, and U. Sattler, "Conjunctive query entailment for shoq," in *Proceedings of the 2007 Description Logic Workshop (DL 2007)*, vol. 250, 2007, pp. 1–11.
14. B. Glimm, C. Lutz, I. Horrocks, and U. Sattler, "Answering conjunctive queries in the $\mathcal{SHIQ}$ description logic," *Journal of Artificial Intelligence Research*, vol. 31, pp. 150–197, 2008.
15. I. Horrocks, U. Sattler, and S. Tobies, "Reasoning with individuals for the description logic shiq," in *Proceedings of the 17th International Conference on Automated Deduction (CADE-17)*, ser. Lecture Notes in Computer Science, D. MacAllester, Ed., no. 1831. Germany: Springer Verlag, 2000, pp. 482–496.
16. I. Horrocks and S. Tessaris, "A conjunctive query language for description logic aboxes," in *In Proc. of the 17th Nat. Conf. on Artificial Intelligence (AAAI-2000)*, 2000, pp. 399–404.
17. U. Hustadt, B. Motik, and U. Sattler, "Data complexity of reasoning in very expressive description logics," in *IJCAI-05, Proc. of the Nineteenth International Joint Conference on Artificial Intelligence, Edinburgh, Scotland, UK, July 30-August 5, 2005*, 2005, pp. 466–471.
18. C. Lutz, "The complexity of conjunctive query answering in expressive description logics," in *Automated Reasoning, 4th International Joint Conference, IJCAR 2008, Sydney, Australia, August 12-15, 2008, Proceedings*, 2008, pp. 179–193.
19. B. Motik and I. Horrocks, "OWL datatypes: Design and implementation," in *Proc. of the 7th Int. Semantic Web Conference (ISWC 2008)*, ser. LNCS, vol. 5318. Springer, October 26–30 2008, pp. 307–322.
20. B. Motik, U. Sattler, and R. Studer, "Query answering for owl-dl with rules," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 3, no. 1, pp. 41–60, 2005.
21. M. Ortiz, D. Calvanese, and T. Eiter, "Characterizing data complexity for conjunctive query answering in expressive description logics," in *Proc. of the 21st Nat. Conf. on Artificial Intelligence (AAAI 2006)*, 2006, pp. 1–6.
22. M. Ortiz, R. Sebastian, and M. Šimkus, "Query answering in the horn fragments of the description logics shoiq and sroiq," in *Proc. of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two*, ser. IJCAI'11. AAAI Press, 2011, pp. 1039–1044.
23. R. Rosati, "On conjunctive query answering in EL," in *Proceedings of the 2007 International Workshop on Description Logic (DL 2007)*. CEUR Electronic Workshop Proceedings, 2007, pp. 1–8.
24. World Wide Web Consortium, "OWL 2 Web ontology language structural specification and functional-style syntax (second edition)."